https://doi.org/10.53656/isct-2025.18

DIFFERENTIAL METHOD FOR SYNTHESIS OF CAM-LEVER MECHANISMS FOR GENERATION OF NON-STRICT MONOTONOUS POSITION FUNCTIONS

Vitan Galabov[1], Roumen Roussev[2], Blagoyka Paleva-Kadiyska[3]
[1)]Technical University, Sofia, Bulgaria
[2)]Faculty of Technics and Technology, Yambol, of Trakia University, Bulgaria
[3)]Todor Kableshkov University of Transport, Sofia, Bulgaria

Abstract. Mechanisms generating periodically variable rotation of the output link (monotonous displacement functions) can be divided into two groups. The mechanisms of the first group generate only strict monotonic functions (without prolonged and even instantaneous dwell of the output link). The second group of mechanisms generates non-strict monotonic functions (with a prolonged or instantaneous dwell of the output link).

This study focuses on functional generator mechanisms with a six-bar topological structure, which are a combination of a cam and lever mechanism.

An original structure of a jointed cam-lever mechanism is proposed, designed to generate strict and non–strict monotonous displacement functions.

The mechanism's kinematic and dynamic characteristics depend on the displacement function type and parameters, or its derivative transfer functions, i.e., on the given law of motion. Laws without a finite or infinite jump (acceleration break) in the second transfer function, or the output acceleration, are preferred. For this purpose, known functions can be used, or functions suitable for the target proposition can be composed.

A law of motion was chosen that is per the requirement for resetting the output velocity, acceleration, and its derivative (jerk) for the boundaries of the output link dwell interval. This is recommended for the synthesis of high-speed cam-lever mechanisms on many technological machines.

A rational mathematical model has been developed, based on elements of differential geometry, for the synthesis of planetary cam-lever mechanisms with a six-bar topological structure and a common geometric axis of rotation of the input and output links.

An example is presented that illustrates the overall process by which a mechanism with the indicated structure was synthesized and which verifies the derived mathematical model.

Keywords: planetary cam-lever mechanisms; law of motion; synthesis

1. Introduction

When designing machines and technical devices with a continuous production process, it is increasingly necessary to synthesize mechanisms for generating monotonic functions for converting uniform rotation into periodical unidirectional uneven rotation when solving several groups of technical problems. One of them is a constant angular velocity drive at the input and temporary dwell of conveyors and conveyor belts, according to the requirements of a certain technological process.

Mechanisms generating rotational periodic rotation of the output link (monotonous displacement functions) can be divided into two groups. The mechanisms of the first group generate only strict monotonic functions (without prolonged and even instantaneous dwell of the output link). Double-crank four-bar linkages and slotted link mechanisms with full rotation of the flat-face (also known as slotted-link) generate approximately specified, strict monotonic displacement functions [1, 2, 3], unlike mechanisms with non-circular gears [4, 5].

The second group of mechanisms generates non-strict monotonic functions (with a prolonged or instantaneous dwell of the output link). This group includes mechanisms with periodical output rotations of the output link in a finite interval of the input rotation (Maltese cross, ratchet, anchor, and other types of mechanisms) [6]. Changes in accelerations with jumps, or inertial load, are avoided by some combined mechanisms involving a cam mechanism [7, 8].

This study focuses on functional generator mechanisms with a six-bar topological structure, which are a combination of a cam and lever mechanism.

The aim is to compile a rational mathematical model, based on elements of differential geometry, for the synthesis of cam-lever mechanisms with a general geometric axis of rotation of the input and output links.

The choice of the law of motion that pursuant with the requirement for resetting the output velocity, acceleration, and its derivative (jerk) for the dwell of the output link dwell interval, which is recommended for the synthesis of high-speed cam-lever mechanisms on many technological machines [9, 10, 11, 12, 13].

2. Structure and mathematical model for mechanism synthesis

An original structure of a jointed cam-lever mechanism is proposed, designed to generate strict and non–strict monotonous displacement functions (Fig.1). The mechanism consists of a jointed parallelogram *OABC* with a double base joint *O*, connecting the input link of length *a*, and the output link of length *b*. Diagonally to the base center is situated the center *B* of a roller with radius *r*, which contacts a fixed cam. For each position, a conditional link can be introduced connecting the center *B* with the corresponding center of curvature B_o of the cam profile c_G, resulting in a six-bar topological structure formed by two sequentially connected four-bar linkages $OABB_0$ and B_0BCO.

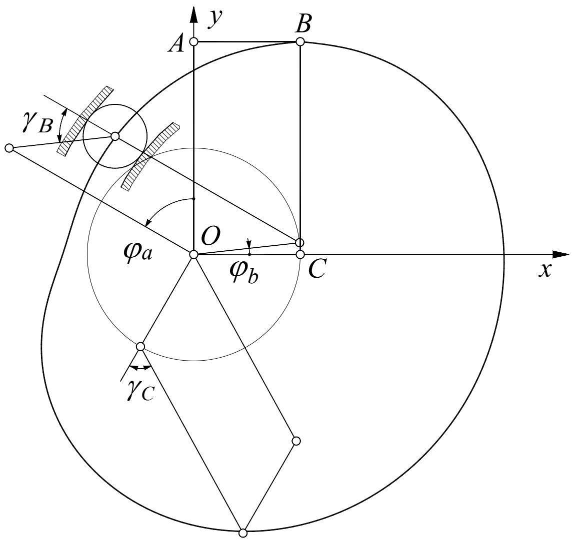


Figure 1. Cam mechanism combined with a jointed parallelogram

The input and output coordinates are the angles φ_a and φ_b relative to the *x*-axis, respectively. In its initial position, the configuration of the jointed parallelogram is rectangular, with the input link *OA* making an angle 90° with the *x* axis. So $\varphi_a = 0.5\pi + \varphi_a^*$, where the initial value of φ_a^* is $\varphi_a^* = 0$. The dependence $\varphi_b = \varphi_b(\varphi_a^*)$ represents the displacement function of the mechanism, and its derivatives $\varphi_b' = d\varphi_b / d\varphi_a$ and $\varphi_b'' = d^2\varphi_b / d\varphi_a^2$ are the first and second transfer functions of the mechanism, respectively. These functions determine the pitch curve c_B of the cam, its profile c_G, and their general evolute, which is the problem of the mechanism synthesis.

The synthesis of the cam profile is performed with a predefined or accepted function of the mechanism displacement $\varphi_b = \varphi_b(\varphi_a^*)$, its derivative transfer functions $\varphi_b' = d\varphi_b / d\varphi_a$, $\varphi_b'' = d^2\varphi_b / d\varphi_a^2$, and the lengths *a* and *b* of the levers of the jointed parallelogram *OABC*.

The pitch curve c_B of the cam is determined by the coordinates of the center *B* as a function of the entrance angle φ_a:

$$(1) \quad \begin{cases} x_B = a\cos\varphi_a + b\cos\varphi_b(\varphi_a); \\ y_B = a\sin\varphi_a + b\sin\varphi_b(\varphi_a). \end{cases}$$

The first derivatives of the coordinates from equation (1) to $\boldsymbol{\varphi_a}$ are necessary to determine the angles of force transmission in the mechanism, and together with their second derivatives, to determine the evolute of the cam profile. For the firsts and seconds derivatives of the coordinates, we get:

$$(2)\quad \begin{cases} x'_B = -a\ \sin\varphi_a - b\varphi'_b \sin\varphi_b(\varphi_a)\ ; \\ y'_B = a\ \cos\varphi_a + b\varphi'_b \cos\varphi_b(\varphi_a)\ ; \end{cases}$$

$$(3)\quad \begin{cases} x''_B = -a\ \cos\varphi_a - b(\varphi''_b \sin\varphi_b + \varphi'^2_b \cos\varphi_b)\ ; \\ y''_B = -a\ \sin\varphi_a + b(\varphi''_b \cos\varphi_b - \varphi'^2_b \sin\varphi_b)\ . \end{cases}$$

The angles of motion transmission γ_B and γ_C, to the centers B and C, respectively, largely determine the force conditions under which the cam-lever mechanism operates. Using well-known formulas from differential geometry, the angular coefficient of the tangent to the pitch curve c_B and the angles of motion transmission are determined:

$$(4)\quad k_t = \frac{y'_B}{x'_B}\ ;$$

$$(5)\quad \gamma_B = \left| arctg \frac{1 + k_b k_t}{k_b - k_t} \right|\ ;$$

$$(6)\quad \gamma_C = \left| arctg \frac{k_a - k_b}{1 + k_a k_b} \right|\ ,$$

where

$$(7)\quad k_a = tg\,\varphi_a\ ; \qquad k_b = tg\,\varphi_b(\varphi_a)$$

The longer the dwell interval, the smaller the values of γ_B and γ_C. This is especially true for γ_B, which can reach unallowable small values.

Therefore, from a hinged parallelogram one can move to a four-bar linkage (Fig. 2), in which the sides *OA* and *OC* retain their lengths *a* and *b*, respectively, while the side *AB* accept a value of $l_{AB} > b$ (for example with 30% to 50%) and for the side *BC* the length $l_{BC} \approx a$ is accept. Thus, the mathematical model for determining the coordinates of the center *B* will take the form:

$$(8)\quad \begin{cases} x_B = a\ \cos\varphi_a - l_{AB}\ \cos(\alpha + \beta + \varphi_a); \\ y_B = a\ \sin\varphi_a - l_{AB}\ \sin(\alpha + \beta + \varphi_a)\ , \end{cases}$$

where

$$(9)\quad \begin{cases} d = \sqrt{a^2 + b^2 - 2\,a\,b\,\cos(\varphi_a - \varphi_b)} \\ \alpha = arc\sin \dfrac{b\,\sin(\varphi_a - \varphi_b)}{d}; \\ \beta = arc\cos \dfrac{l_{AB}^2 - l_{BC}^2 + d^2)}{2\,l_{AB}\,d}\ ; \end{cases}$$

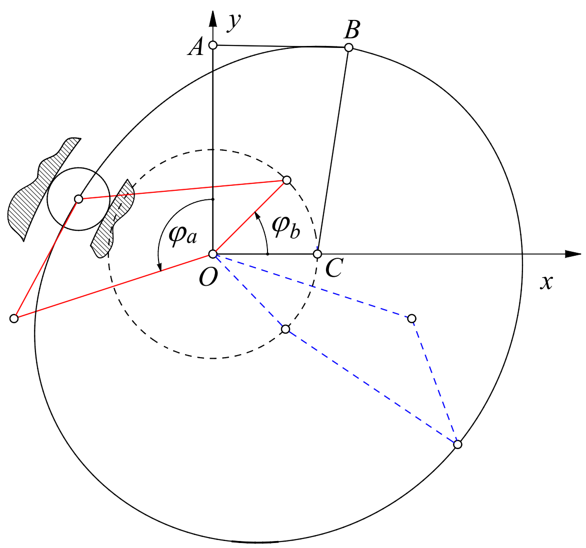


Figure 2. Cam mechanism combined with a four-bar linkage

The cam profile c_G is the equidistant of the pitch curve c_B, located at a distance r (the radius of the roller) along the normal n with an angular coefficient $k_n = -(k_t)^{-1}$:

$$(10)\quad x_G = x_B \pm \frac{r.\mathrm{si\,gn}\,k_n}{\sqrt{1+k_n^2}}; \qquad y_G = y_B \pm \frac{r.k_n\,\mathrm{si\,gn}\,k_n}{\sqrt{1+k_n^2}}.$$

The "+" and "-" signs in equations (8) correspond to the outer and inner cam profiles relative to the roller. The curve of center of a machining tool can be determined from equations (8) by replacing r with r - r_i, where r_i is the radius of the machining tool.

The radius of curvature of the curve of center ρ_B and the cam profile ρ_G are determined by a well-known formula from differential geometry:

$$(11)\quad \rho_B = \frac{\sqrt{(x_B'^2 + y_B'^2)^3}}{x_B' y_B'' + y_B' x_B''}; \qquad \rho_G = \rho_B \pm r \,.$$

Their common center of curvature

$$(12)\quad \begin{cases} x_{B_0} = x_B - \dfrac{y_B'(x_B'^2 + y_B'^2)}{x_B' y_B'' + y_B' x_B''}, \\ y_{B_0} = y_B + \dfrac{x_B'(x_B'^2 + y_B'^2)}{x_B' y_B'' + y_B' x_B''} \end{cases}$$

determines their common evolute ε when x_A, y_B and their derivatives are functions of the angle φ_a.

3. Choice of the law of motion of the output link

The mechanism's kinematic and dynamic characteristics depend on the displacement function type and parameters, or its derivative transfer functions, i.e., on the given law of motion. Laws without a finite or infinite jump (acceleration break) in the second transfer function, or the output acceleration, are preferred. For this purpose, known functions can be used, or functions suitable for the target proposition can be composed.

According to the technological process, the intervals of uniform rotation of the input link and dwell of the output link 3 are determined. In the range of motion of the output link, the engineer must select an appropriate output rotation law. It is possible to choose an output rotation function with high-order osculation towards the *x*-axis in the entire interval of the input coordinate φ_a, which practically leads to a prolonged dwell of the output link. In the article, the authors present this possibility.

The displacement function of the output link $\varphi_b = \varphi_b(\varphi_a^*)$ and its derivative transfer functions $\varphi_b' = d\varphi_b / d\varphi_a$, $\varphi_b'' = d^2\varphi_b / d\varphi_a^2$, $\varphi_b''' = d^3\varphi_b / d\varphi_a^3$ are conveniently described by the normalized function $u(\xi)$ and its derivatives

$$(13)\quad \begin{cases} \varphi_b = 2\pi \,.\, u(\xi), \\ \varphi_b' = u'(\xi), \\ \varphi_b'' = \dfrac{1}{2\pi} u''(\xi), \\ \varphi_b''' = \dfrac{1}{4\pi^2} u'''(\xi), \end{cases}$$

Shock-free operation of the mechanism can be achieved by means of polynomial $u(\xi) \in [0,\ 0.5]$ of exponent 4-5-6-7 in the interval $\varphi_a^* \in [0,\ \pi]$, resp. $\xi \in [0,\ 0.5]$. Such a polynomial and its derivatives are shown below

$$(14)\quad \begin{cases} u(\xi) = 35\xi^4 - 84\xi^5 + 70\xi^6 - 20\xi^7, \\ u'(\xi) = 140(\xi^3 - 3\xi^4 + 3\xi^5 - \xi^6), \\ u''(\xi) = 420(\xi^2 - 4\xi^3 + 5\xi^4 - 2\xi^5), \\ u'''(\xi) = 840(\xi - 6\xi^2 + 10\xi^3 - 5\xi^4) \end{cases}$$

which are taking a value of zero until the third derivative at $\xi = 0$, respectively $\varphi_a^* = 0$. The following extremum values are obtained $u'_{\max} = 2.188$, $u''_{\max} = 7.513$, $u''_{\min} = -7.513$, $u'''_{\max} = 42$, $u'''_{\min} = -52.5$, respectively to $u'(\xi)$, $u''(\xi)$ and $u'''(\xi)$.

Fig. 3 presents the functions (13) at $\varphi_a^* \in [0,\ \pi]$, and in the interval $\xi \in [0, -0.5]$

respectively $\varphi_a^* \in [0,\ -\pi]$ the following relations are used: $\underline{\varphi}_b = -\varphi_b$; $\underline{\varphi}_b{}' = \varphi_b{}'$; $\underline{\varphi}_b{}'' = -\varphi_b{}''$; $\underline{\varphi}_b{}''' = \varphi_b{}'''$.

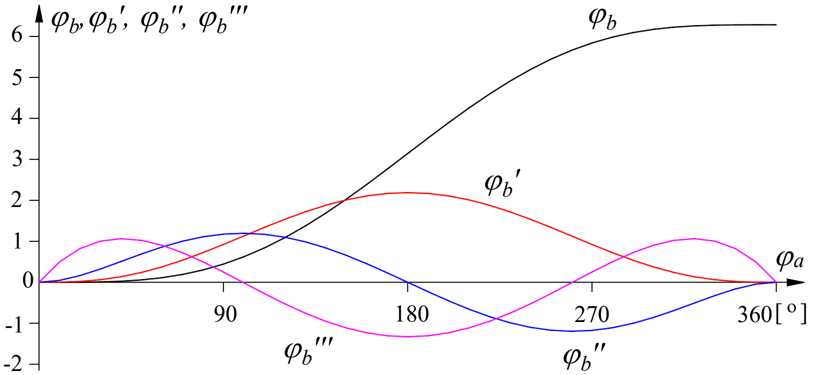


Figure 3. Displacement function and derived transfer functions

Example. To synthesize a mechanism of the considered type for driving a conveyor with appropriately chosen values a = 200 mm, b = 100 mm, r = 30 mm.

From equations (11) and a choice normalized polynomial and its derivatives (14), the displacement function of the output link $\varphi_b = \varphi_b(\varphi_a^*)$ and its derivative transfer functions $\varphi_b' = d\varphi_b / d\varphi_a$, $\varphi_b'' = d^2\varphi_b / d\varphi_a^2$, $\varphi_b''' = d^3\varphi_b / d\varphi_a^3$ are determined (Fig. 2).

From equations (1) the pitch curve c_B is determined (Fig. 1). From equations (2) and (3) the derivatives x_B', y_B', x_B'', y_B'' are determined. After determining from (4) the angular coefficient k_t of the tangent to the pitch curve c_B from (5), (6) and (7) the angles of motion transmission γ_B and γ_C are determined (Fig. 4). The synthesis ends with the determination of the cam profile c_G and its evolute ε according to equations (10) and (12), respectively.

The kinematic diagram of the synthesized mechanism is shown in Fig. 1.

The dwell interval is $\varphi_{dwell} = 10°$ at a deviation of $|\Delta\varphi_b| = 1.10^{-5}$ and $\varphi_{dwell} = 15°$ at a deviation of $|\Delta\varphi_b| = 1.10^{-4}$. The dwell time is approximately 1/36 and 1/24 of the entire geometric period $\varphi_T = 2\pi(360°)$ of the mechanism, respectively.

Graphs of the motion transmission angles determined by equations from (4) to (7) are presented in Fig. 4.

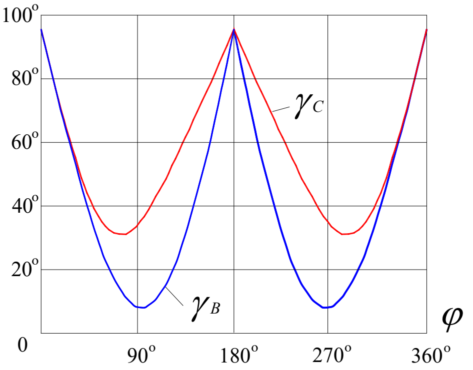


Figure 4. The angles of motion transmission γ_B and γ_C

Conclusion

This study focuses on functional generator mechanisms with a six-bar topological structure, which are a combination of a cam and lever mechanism. The lengths of its links are determined by requirements for maximum sizes, according to the purpose of the mechanism, as well as by the influence of the *b/a* ratio on the curvature the curve of center of the cam. The smaller the ratio b/a, the smaller the variation in the curvature the curve of center. But the reduction in length b is limited by design considerations related to the sizing of the bearing assemblies of the mechanism and the radius r of the roller.

A law of motion was chosen that is per the requirement for resetting the output velocity, acceleration, and its derivative (jerk) for the boundaries of the output link dwell interval. A rational mathematical model has been developed, based on elements of differential geometry, for the synthesis of planetary cam-lever mechanisms with a six-bar topological structure and a common geometric axis of rotation of the input and output links. This makes it possible to fully reveal the geometry of the cam mechanism, the pitch curve of the roller and of various machining tools, the angles of motion transmission, and the common evolute of the cam profile.

The presented example illustrates the overall process by which mechanisms with the indicated structure are synthesized and proves the correctness of the developed mathematical model.

REFERENCES

[1] SUH C. H., RADCLIFFE C. W., Kinematics and Mechanisms Design. New York: John Wiley & Sons, 1978, 434 p. ISBN 978-0-471-01461-4.

[2] PEISAKH E. E., NESTOROV V. A., System of designing flat lever mechanisms. Moscow, Mashinostroenie, 1988, p. 232. (in Russian).

[3] ERDMAN A. G., SANDOR G. N., Mechanism Design: Analysis and Synthesis. 2nd Edition. New Jersey: Prentice-Hall Inc., 1991, vol. 1. 650 p. ISBN 978-0-13-573536-7.

[4] LITVIN F. L., Theory of Gearing. Moscow, Nauka, 1968, p. 584. (in Russian).

[5] GALABOV V., On the synthesis, analysis, and production of non-circular gears, Dissertation for Candidate Doctor of Science, Higher Mechanical and Electrical Engineering Institute-Sofia, 1975, p. 222.

[6] KOZHEVNIKOV S. N., ESIPENKO YA. I., RASKIN YA. M, Mechanisms (handbook), Moscow, Mashinostroenie, 1976, p. 784. (in Russian).

[7] KARELIN V. S, Design of lever and gear-lever mechanisms (handbook), Moscow, Mashinostroeniya, 1986, p. 181. (in Russian).

[8] VOLMER J., Getriebetechnik – Kurvengetriebe – CAD/CAM, Verlag Technik, Berlin, 1989 Congress on TMM, pp. 1303-1306.

[9] ROTHBART, H., A., CAMES. Design, Dynamics and Accuracy, N., Y., John Wiley & Sons,1965, p. 336. [10] SADEK, K. S. H., DAADBIN, A., Improved Cam Profiles for High-Speed Machinery Using Polynomial Curve Fitting, J. of Process Mechanical Engineering, Vol.204, n.E 2, 1990, pp.127 – 132.

[11] BLECHSMIDT, J. L., LEE, C. H., Design and Analysis of Cam Profiles Using Algebraic Functions, ASME, DE, vol. 32 pt. 2, New York, pp. 451 – 459, 1991.

[12] NORTON R., L., Design of Machinery, McGraw-Hill Inc., New York, 306 p., 1992, ISBN 978-0-07-047808-4.

[13] CARDONA A., LENS E. and Nigro N., Optimal Design of Cams, Multibody System Dynamics, vol. 7, no. 3, pp. 285 – 305, 2002, ISSN 1384-5640. DOI 10.1023/A:1015243126749.

✉ **Prof. Vitan Galabov, DSc.**
ORCID iD: 0009-0001-1527-6099
Technical University-Sofia
8, Kl. Ohridski Blvd.1000 Sofia, Bulgaria
E-mail: vgalabov@abv.bg

✉ **Dr. Roumen Roussev, Assoc. Prof.**
ORCID iD: 0009-0001-8485-1229
Faculty of Technics and Technology, Yambol, Trakia University
38, Graf Ignatiev St., 8600 Yambol, Bulgaria
E-mail: roussev_r@abv.bg

✉ **Dr. Blagoyka Paleva-Kadiyska, Assoc. Prof.**
ORCID iD: 0000-0002-8514-4542
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: paleva-kadiyska.bl@abv.bg; bip-kadiyska@vtu.bg

https://doi.org/10.53656/isct-2025.19

Mechanics, Mechatronics and Mathematics

TYPE SYNTHESIS OF GUIDANCE MECHANISMS WITH A FOUR-BAR TOPOLOGICAL STRUCTURE

Blagoyka Paleva-Kadiyska[1], Vitan Galabov[2]
[1)]Todor Kableshkov University of Transport, Sofia, Bulgaria
[2)]Technical University, Sofia, Bulgaria

Abstract. The functional approach (method) for structural-dimensional synthesis of mechanisms has been applied for type synthesis of guidance mechanisms with a four-bar topological structure. According to the method, their typological structures are synthesized by superimposing a three-bar conditioned open kinematic chain, including a base (fixed link), an initial and an end link, with a control functional generator mechanism of a different type, kinematically connecting the base, with the end link of the chain at a relative stand the initial link.

The type synthesis also includes a possible replacement of some links and the kinematic lower pairs assigned to them in the structure of the mechanism, with profiled kinematic higher pairs. This completely determines the type of mechanism. In this sense, a *topological structure* of a mechanism with kinematic higher pairs means that these pairs are reduced to kinematic lower pairs. Thus, guidance cam mechanisms with two moving links, after kinematically replacing the cam pair with two kinematic lower pairs, have a four-bar topological structure, as do the mechanisms analysed in this article.

The purpose of this study is to apply the functional method for structural-dimensional synthesis of mechanisms for type synthesis of guidance planar mechanisms with a four-bar topological structure, including a three-bar conditionally open kinematic chain and a control functional generator mechanism of a different type.

A type systematics of guidance planar mechanisms with a four-bar topological structure was carried out in two stages: determination of the types of three-bar conditionally open kinematic chains; determination of the types of control functional generator mechanisms. Guidelines/Directives are given for the preliminary and final determination of the type structure of a guidance mechanism before the engineer proceeds to its dimensional synthesis.

Keywords: type synthesis; topological synthesis; cam mechanisms

1. Introduction

The terminology used in publications related to the synthesis of mechanisms is not unambiguously stated (unalterable). Researchers [1, 2], mostly European, divide the synthesis of mechanisms into two stages: 1) *structural synthesis*, including the selection of a structural diagram of the mechanism, which clarifies the type of mechanism, the number of links and kinematic pairs, their type, and alignment; 2) *kinematic synthesis*, in which the geometric parameters of the mechanism diagram in its project position are determined according to given geometric and kinematic conditions.

Sandor and Erdman [3] divide the field of mechanism synthesis into two categories: 1) *type synthesis*, in which a presumed type of mechanism is determined (lever, cam, gear, or other type) and includes the subcategory of *integer* (*Diophantine*, *number*) *synthesis* [4, 5]; 2) *dimensional* (*dimensioning*) *synthesis*, in which the mechanism diagram is dimensioned [6]. The terms *dimensional* and *kinematic synthesis* are often seen as synonyms [7]. Kinematic synthesis is often complemented by *dynamic synthesis* under given dynamic conditions.

In integer synthesis, the configurations of mechanisms are defined topological, but not dimensional. Therefore *integer* and *topological synthesis* are perceived as synonyms. In *topological synthesis*, *topological structures* of mechanisms are built, determined by the number of their links, the number of their kinematic lower pairs (without distinguishing between rotating and sliding), and their alignment [6, 7, 8]. These structures are represented as non-dimensional lever mechanisms with only kinematic lower pairs.

Methods for generating topological kinematic structures of mechanisms can be divided into four types [9], based on: – the groups of Assur [1]; – closed kinematic chains in Franke and Gruebler notation [10]; – closed kinematic chains in graph notation [11]; – combinations of the above methods [12].

Type synthesis is usually identified with the systematization (classification) of mechanisms. Instead of *type synthesis*, the term *typological synthesis* is often used. Most authors consider *topological synthesis* to be the first phase of the typological synthesis of mechanisms. In type synthesis, typological structures of mechanisms are built, determined by the number of their links, the number and type of their kinematic pairs (lower and higher), and their arrangement [6, 8].

The type synthesis also includes a possible replacement of some links and the kinematic lower pairs assigned to them in the structure of the mechanism, with profiled kinematic higher pairs. This completely determines the type of mechanism. In this sense, a *topological structure* of a mechanism with kinematic higher pairs means that these pairs are reduced to kinematic lower pairs. Thus, guidance cam mechanisms with two moving links, after kinematically replacing the cam pair with two kinematic lower pairs, have a four-bar topological structure, as do the mechanisms analysed in this article.

In Bulgarian technical terminology, the concepts of *kinematic* (*structural* and *dimensional*) and *dynamic synthesis* have been adopted. The terms topological and typological synthesis also appear in scientific *publications, as well as in this article, as stages of structural synthesis.*

After a critical analysis of the methods for structural synthesis, Galabov [9] proposes a functional approach for the structural-dimansional synthesis of mechanisms. According to the method, the mechanisms are built from interconnected conditionally open kinematic chains (OKC) with different functional purposes, in which they form closed chains of control functional generator mechanisms (CFGM), the frame (base) of which are movable links of the chains. According to the method, the typological structure of elementary guidance mechanisms is formed by superimposing a three-bar conditionally open kinematic chain, consisting of a frame (base), an initial link and an end link, connected in series with kinematic lower pairs, with a control functional generator mechanisms of a different type, which kinematically connects the frame with the end link, respectively, as the input and output link of the control functional generator mechanism, with a relative base the initial link.

The purpose of this study is to apply the functional method for structural-dimensional synthesis of mechanisms for type synthesis of guidance planar mechanisms with a four-bar topological structure, including a three-bar conditionally OKC and a CFGM of a different type.

2. Type classification of planar lever mechanisms with a four-bar topological structure

To determine the structures of the anisomorphic functional generator and guidance four-bar mechanisms, according to the now classic approach for structural synthesis of mechanisms based on closed loops, we start from a *structural graph* of a quadrilateral closed kinematic chain (Fig.1 – first row). Lines indicate the links, and circles indicate the kinematic lower pairs, without distinguishing between rotating (R-pairs) and sliding (P-pairs). This graph illustrates the *topological structure* of four-bar lever mechanisms. By introducing all possible combinations without repetitions of rotating (**O**) and sliding (**◻**) kinematic pairs, four anisomorphic type structures (from I to IV) of four-bar closed kinematic chains are obtained (Fig.1 – second row).

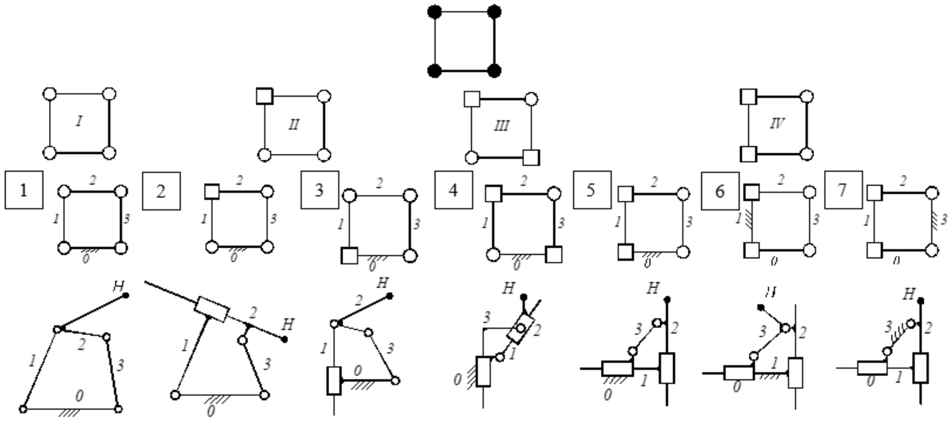


Figure 1. Anisomorphic structures of four-bar mechanisms

From these chains, with different choices of the base (link assumed to be fixed), the type structures of seven types of four-bar mechanisms are obtained (Fig.1 – third row). With the initial and final link (links 1 and 3) kinematically connected to the fixed link 0, the first 5 type structures of functional generator four-bar lever linkages are obtained. The next two structures (5 and 6) have bases, respectively, links 1 and 3. Next to each structure, in the last row of Fig. 1, a kinematic diagram (dimensioned type structure after dimensional synthesis) of a linkage functional generator mechanism is given: 1) four-bar mechanism; 2) slotted link mechanism; 3) slider-crank mechanism; 4) tangent mechanism; 5) sinusoidal mechanism; 6) ellipsograph mechanism; 7) Oldham coupling mechanism.

Guidance four-bar mechanisms are obtained if the executive links are those with a common planar movement (links with number 2 for diagrams 1 to 5), which, with their so-called characteristic point ***H*** describe desired trajectories. For the mechanisms of diagrams 6 and 7, the units with common planar motion are 3 and 1, respectively. The trajectories of all points of the sinusoidal mechanism (diagram 5) are circles with a radius equal to the length of the crank 3, since the link 2 performs a circular translation. The trajectories of the points on the plane of link 3 of the ellipsograph mechanism (diagram 6) are ellipses.

3. Type classification of planar guidance mechanisms with a four-bar topological structure

According to the functional method for structural-dimensional synthesis of mechanisms [9] the type structure of the considered class of mechanisms is synthesized by superimposing a three-bar conditionally OKC, consisting of base 0, initial link 1 and end link 2, connected in series with kinematic lower pairs (Fig.2, first row), with a CFGM, which kinematically connects base 0 with end link 2, respectively, as input and output links of the OKC with relative base initial link 1.

The type synthesis of these mechanisms involves two stages.

Stage 1 – Determination of the types of three-bar conditional OKC

Four anisomorphic structures of three-bar OKC are possible, which can be denoted by the symbols of the kinematic pairs between the links: RR, RP, PR and PP. The chains have two degrees of freedom, and if they remain open, two controlled motors are required, one for each moving link. The choice of the OKC structure depends on the specific problem and the possibilities for structural decoupling of motion. Chains with RR structures have a universal purpose. They are suitable for mechanisms designed to generate trajectories close to an arc of a circle, cycloid, epicycloid, hypocycloid, ellipse, straight line segment, polygons with rounded vertices or with cuspate points, etc.. Considering also the favorable design and performance characteristics of R-pairs, OKC with RR-structure are preferable except in cases where the trajectory is close to rectilinear with a relatively large stroke. In such cases, chains with PR-structure usually have the advantage.

For helical trajectories, chains with RP-structure are preferable. This structure, with length $l_1 = 0$ of initial link 1, is suitable for curves described in polar coordinates. It allows for decoupling (independence) of the equal rotation of links 1 and 2 from the rectilinear translation of link 2. But in this case, two CFGMs are needed. Each of them, independently of the other, changes one of the polar coordinates of the characteristic point *H*, describing the target trajectory [9].

The PP-structure is also suitable for curves described in Cartesian coordinates. It allows for decoupling of the rectilinear translations of link 1 relative to frame (base) link 0 and of link 2 relative to link 1. And in this case, two CFGMs are needed, each of which, independently of the other, changes one of the Cartesian coordinates of point *H*. In both decoupled of OKC structures, it is possible to drive two cams and links 1 and 2, respectively, with one motor and functional generator mechanism (for example, a gear). But in this case, the mechanism no longer has a four-bar topological structure.

A precise choice of one of the selected competitive structures can be made after dimensional the OKC, determining the transfer functions of the CFGM, and their kinematic diagrams. These are tasks of the dimensional synthesis of this class of mechanisms that remain to be solved.

Stage 2 – Determination of the types CFGM

The elementary way to close a three-bar OKC is to connect the fixed link 0 with the end link 2 with a binary link 3 (the last column in Fig.1). A guidince *four-bar linkage* is formed, in which a linkage CFGM is built. The structure of these elementary linkages can be considered *topological* if the dimensions and the difference between the rotating and sliding pairs are ignored.

Different types of guidance mechanisms with a four-bar topological structure are presented in four columns in Fig.2 with RR, RP, PR, and PP structures of conditional OKC, respectively. CFGM on the second row in Fig.2 (diagrams 1, 2, 3, and 4) is *cam* *mechanism* with an input cam fixed to the base 0 and an end link 2 with a roller. By changing the positions of the cam and the roller, a mechanism with an output profiled link is obtained (inverse cam mechanism).

CFGM on the third row in Fig.2 (diagrams 5, 6, 7, and 8) has two cams. The input link is the cam fixed to the frame 0, and the other cam is fixed to the end link 2. If the cams are formed along the *centroids of the relative motion* with respect to the relative position 1 of the CFGM, then the is a *centrode mechanism*. If the profile of the cam contacting link 2 is rectilineal, a control cam mechanism with a flat output link will result – a plate slider or a flat-face. When the positions of the cam and slotted link are changed, the profile of the cam towards base 0 is rectilineal. An *inverted cam mechanism* is obtained.

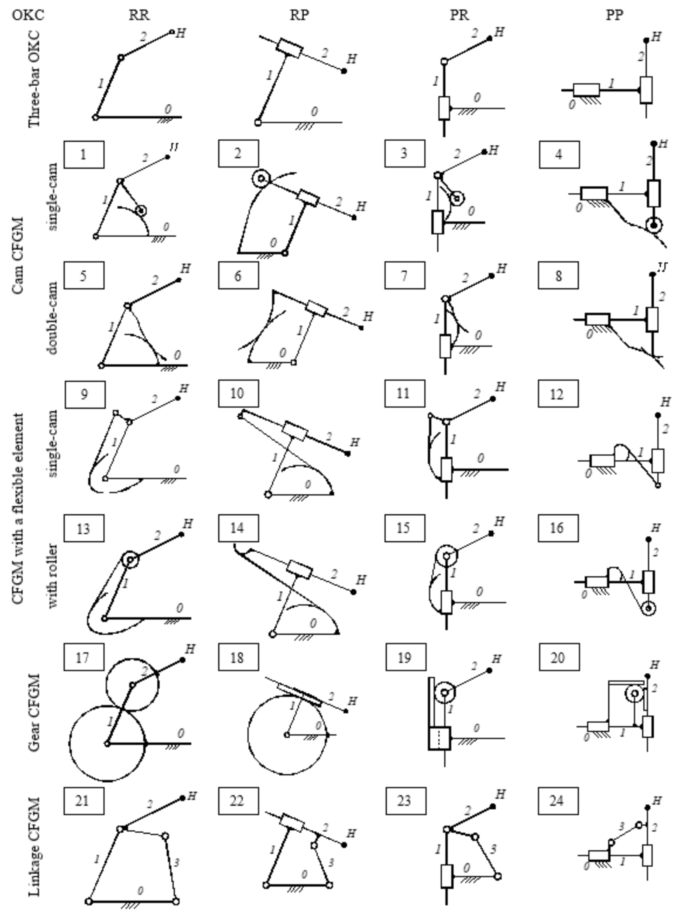


Figure 2. Anisomorphic three-bar OKC and guidance mechanisms with a four-bar topological structure

The CFGM in the fourth row in Fig.2 (diagrams 9, 10, 11, and 12) are *mechanisms with a flexible inextensible element* (tape, cord, timing belt, multi-strand steel rope, etc.), hingedly connected to the end link 2. If the cam is fixed to link 2, and the flexible link is joined to frame 0, then an *inverse mechanism with a flexible element* (*band mechanism*) is obtained.

The CFGM on the fifth row in Fig.2 (diagrams 13, 14, 15 and 16) is a *band mechanism* with a roller end link 2. By changing the positions of the cam and the roller, an inverse mechanism with a flexible element with roller input link 0 is obtained. If the roller is non-circular, a *double-cam band mechanism* is obtained.

The CFGM of the sixth row in Fig.2 (diagrams 17, 18, 19, and 20) is a gear mechanism with two circular gears (diagram 17) or with non-circular gears, with a gear and a rack (diagrams 18 and 19) or a gear and two racks (diagram 20).

The possibilities for generating different trajectories with guidance mechanisms are unusually large, even with CFGMs with a constant transfer ratio. For example, with different transfer ratios i of the CFGM and ratios $\alpha = l_1/l_2$ of the lengths l_1 and l_2 of links 1 and 2, 84 typical trajectories are obtained, of which: • $\lambda = 1$ and $i > 1$ *the curves of Gvido-Grandi,* of which at $i = 4$ *four-leaf clover;* • $\lambda = 2$ and $i = 3$ *the curves of Steiner*; • $\lambda = 3$ and $i = 4$ the asteroid; • $i = 1/2$ *Pascal's snails, of which at* $\lambda = 1/2$ the cardioid; • $i = 2$ different *ellipses*, at $\lambda = 1$ a *straight line segment*, and at $l_2 = 0$ a circle, etc.

Curves with *cuspate points* are obtained at $\lambda = i - 1$, and at $\lambda = -(i - 1)^2$ – curves with flat (*rectifiable*) points. The geometric meaning of the sign (–) means that *flat points* are obtained when the links of a conditional OKC are folded.

CFGM of the sixth row in Fig. 2 (diagrams 21, 22, 23, and 24) are four-bar mechanisms. They can be considered as particular cases of functional generator mechanisms with a kinematic higher pair, for example, a cam mechanism with a cam profiled along an arc of a circle. They can also be considered as particular cases of *band mechanisms*, in which the cams are folded into points (joints) and the flexible element is replaced by a connecting lever. If the two points are extremely distant, then they define the axes of two R-pairs of the connecting lever, as can be seen from diagrams 21, 22, 23, and 24. In the RR-structure of a conditional OKC, the following is obtained: *a four-bar mechanism* (diagram 21); in the RP-structure – *a slider mechanism* with one slider (diagram 22); in the PR-structure – *a crank-slider mechanism* (scheme 23); in the PP-structure - *an ellipsographic mechanism* (diagram 24).

With an infinitely distant center of rotation between links 0 and 3, a second slider appears in the slotted lever mechanism of diagram 22. *A slotted lever mechanism with two sliders* (in particular, *tangential*) is obtained, located between joints (Fig.1, diagram 4). When changing the base (from link 0 to link 1) of the ellipsographic mechanism (diagram 24), a *sinusoidal mechanism* is obtained (Fig.1, diagram 5). If the base is link 3, then an *Oldham coupling mechanism* is obtained (Fig.1, diagram 6) with a gear ratio of 1.

The choice of CFGM depends on the specific problem and the possibilities for structural decoupling of motions. If there is no requirement for the accurate generation of a given trajectory, it is logical to seek a solution with a lever CFGM. The engineer may choose such a choice of CFGM if a stable, relatively large-sized mechanism is required, dictated by target trajectories with relatively large sizes. Examples of this are the main mechanisms of specialized robots, heavily loaded manipulation mechanisms of lifting and handling machines, etc.

On the other hand, lever CFGMs are usually preferred as micro mechanisms and even grippers, due to their relatively smaller mass compared to other types of CFGMs. Driving the grippers with a linear motor predetermines the introduction of a slider as an input link and a crank-slider CFGM. When synthesizing such and other guiding lever four-bar mechanisms, for a given input-output force ratio, it is possible that the intermediate joint between the crank and the connecting rod on the CFGM may be unacceptably distant from a technical viewpoint. Then it is appropriate to replace this joint with a slider, and then perform dimensional synthesis again. The probability of obtaining a successful technical solution is now significantly higher.

Engineers most often turn to cam-type CFGMs for incorporation into structures of strictly guidance CFGMs with relatively small sizes. Typical examples are magnetic tape memory mechanisms in storage devices. The same applies to gear CFGMs of robots in microtechnology, with the exception of those with gear racks, when they are used for translations over significant distances, for example, in some manipulators and metal-cutting machines.

Various manipulation mechanisms of specialized robots and lifting and handling machines incorporate CFGMs with flexible links, as they enable them to transmit movements and forces over significant distances. A number of applications of these CFGMs are also found in various devices and robots, including robots in microtechnology.

The question of choosing the type of CFGM remains open in cases where several types of CFGM, determined after preliminary selection, successfully compete as the first stage of the typodimensional synthesis method. According to the method, a joint dimensional synthesis of competing types of mechanisms with a common mathematical model is performed. The prospective kinematic diagrams are determined, from which a final choice is made by type and size of the CFGM according to an objective function, which is composed of quantitative estimates and weight coefficients of selected optimization criteria. But this is not the subject of this article.

Conclusion

The *functional method for structural-dimensional synthesis of mechanisms* for their type synthesis of the considered class of guidance mechanisms with a four-bar topological structure has been applied. According to the method, the type structure of elementary guidance mechanisms is synthesized from the stratification (lamination) of a three-bar conditionally OKC. The kinematic chain is composed of a fixed link, an initial link, and an end link. They are connected in series with kinematic lower pairs, with a CFGM of a different type, which kinematically connects the base with the end link, respectively, as the input and output link of the CFGM with a relative fixed link, the initial link.

A type classification of planar guidance mechanisms with a four-bar topological structure has been done. Two stages of type-dimensional synthesis are distinguished: (1) determination of the types of three-bar conditionally OKC; (2) determination of the types of CFGMs.

Recommendations are given for preliminary and final determination of the typological structure of a conditionally OKC and the type of the CFGM. This guides the engineer to the most appropriate typological structure of a guidance mechanism according to the technical problem set, before proceeding to its dimensional synthesis.

REFERENCES

[1] ARTOBOLEVSKY I. I., LEVITSKYI N. I., CHERKUDINOV S.A. Synthesis of plane mechanisms. Moscow, Fizmatgiz, 1084 p., 1959 (in Russian).

[2] LICHTENHELDT, W. , LUCK, K., Konstruktionslehre der Getriebe, Akad.-Verlag, Berlin, 354 p., 1979.

[3] ERDMAN, A. G., SANDOR, G. N., Mechanism Design: Analysis and Synthesis, Prentice-Hall Inc., New Jersey, Vol. 1, 1984, 2-nd Edition, 1991, ISBN 978-0-13-573536-7.

[4] JOHNSON, R. C., TOWFIGH, K., Application of Number Synthesis to Practical Problems in Creative Design , ASME Paper No. 65 -WA / MD -9, 1965.

[5] TISCHLER, C. R., SAMUEL, A. E., HUNT, H. K., Kinematic Chains for Robot Hands – I Orderly Number Synthesis, II - Kinematics Constrains, Classification, Connectivity and Actuation, Mech. Mach. Theory, Vol. 30, No 8, pp. 1193-123, 1995, ISSN 0094-114X. DOI 10.1016/0094-114X(95)00045-8.

[6] ERDMAN, A. G., (editor) Modern Kinematics, John-Wiley & Sons, Inc. New York, 604 p., 1993, ISBN 978-0-471-55459-2.

[7] TERMINOLOGY for the Theory of Machines and Mechanisms, Mech. Mach. Theory, Vol. 26, No. 5, pp. 453-539, 1991, ISSN 0094-114X, doi 10.1016/0094-114X(91)90054-2.

[8] OLSON, D.G., RILEY D. R., ERDMAN, A. G., A Systematic Procedure for Type Syntsesis of Mechanisms with Literature Revew, Mechanism & Macine Theory, Vol. 20 (4), pp. 285-295, 1985, ISSN 0094-114X, doi 10.1016/0094-114X(85)90032-1.

[9] GALABOV V., Structural & Dimensional Synthesis of Mechanisms, DSc-technical sciences dissertation, TU – Sofia, 491 p., 1998. (in Bulgarian).

[10] CROSSLEY, F. R. E., A Contribution to Gruebler's Theory in the Number Synthesis of Planar Mechanisms, Trans. ASME, J. of Eng. for Industry, Ser.B, Vol. 86, No 1, pp. 1 – 8, 1964, ISSN 0022-0817, doi 10.1115/1.3670494.
[11] CROSSLEY F. R. E., The Permutation of Kinematic Chains of Eight Members or Less from the Graph Theoretic Viewpoint, Developments in Theoretical and Applied Mechanics, Pergamon Press, Oxford, Vol. 2, pp. 467 – 486, 1965.
[12] KONSTANTINOV, M., Strukture der Mechanismen und Roboter, Wissenschaftliche Zeitschrift der Technischen Universitat Dresden, 38, Heft 5/6, pp. 159-167, 1989.

✉ **Prof. Vitan Galabov, DSc.**
ORCID iD: 0009-0001-1527-6099
Technical University-Sofia
8, Kl. Ohridski Blvd., 1000 Sofia, Bulgaria
E-mail: vgalabov@abv.bg

✉ **Dr. Blagoyka Paleva-Kadiyska, Assoc. Prof.**
ORCID iD: 0000-0002-8514-4542
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: paleva-kadiyska.bl@abv.bg; bip-kadiyska@vtu.bg

https://doi.org/10.53656/isct-2025.20

EXPERIMENTAL STUDY AND ANALYSIS OF THE ENERGY PERFORMANCE OF THYRISTOR RECTIFIERS FOR TRACTION SUBSTATIONS

Todor Lalev, Martina Tomcheva, Detelina Pavlova
Todor Kableshkov University of Transport, Sofia, Bulgaria

Abstract. This paper presents the results of an experimental study on the energy performance of thyristor rectifiers used in traction substations for supplying trolleybus and tram electric transport. The purpose of the study is to assess the efficiency of thyristor traction units that were introduced in the 1990s and are still part of the urban electric transport infrastructure in Sofia.

The paper includes results from experimental measurements of key electrical parameters under various operating modes of thyristor rectifier units. The study was conducted in real operational conditions in a traction substation equipped with such devices. The measured parameters include input values such as voltage, current, ripple, harmonic distortions, power factor, cos φ, as well as active and reactive energy.

Power losses, current ripple levels, and the behavior of the rectifiers under different operating regimes were analyzed. The obtained results provide a basis for a well-founded evaluation of the technical suitability of the systems in relation to modern requirements and for formulating guidelines for improving their energy efficiency. This paper contributes to a deeper understanding of the operational characteristics of thyristor rectifiers and their role in the sustainable functioning of the electric transport network.

Keywords: rectifier substation; thyristor rectifier; energy efficiency

1. Introduction

The municipal company Sofia Public Electrical Transport Company JSC is the primary operator of electric urban transport in Sofia, the capital of Bulgaria. With over 100 years of history, its infrastructure utilizes various types and generations of Traction rectifier stations (TRS), used to supply power to the trolleybus and tram networks. Until 2012, the company played a key role as the backbone of urban transport in Sofia, before the significant expansion of the Sofia metro. Despite the changes in the transportation structure, electric transport still plays a crucial role in the city's mobility. After the introduction and expansion of the metro, the importance of trams and trolleybuses has not diminished – on the contrary, they remain an indispensable element of connectivity between metro lines and ensure transportation accessibility in areas where the metro does not reach. Due to their lower construction and operational costs, as well as greater flexibility in route changes, electric transport continues to be an effective and eco-friendly alternative, easily adaptable to future developments in the urban transport network.

In the context of Bulgaria's membership in the European Union, sustainability, energy efficiency, and environmental standards are leading priorities in the development of transport infrastructure [2]. This is where electric transport plays a strategic role in building a sustainable and interconnected urban transport system.

Despite various challenges – financial, organizational, and technological – Sofia Public Electrical Transport Company JSC has succeeded in not only modernizing its infrastructure but also maintaining older facilities in excellent technical condition. Thanks to this, its network includes and operates different generations of technologies, which provide an excellent foundation for conducting experimental research and practical analysis. This includes a variety of technical solutions related to rectifier units and traction power supply systems, which are still used in a real operational environment.

2. General

The subject of study in this paper is the rectifier substation Hadji Dimitar, located in the northeastern part of Sofia. The station is equipped with thyristor rectifiers, making it a highly interesting object for analysis regarding its performance under real operational conditions. This is why it was selected for the study, focusing on the energy characteristics of TRUs operating in TRS.

2.1. Description of the object of study

The Hadji Dimitar TRS was commissioned in 1995. It is located in the eastern part of Sofia and serves the electric urban transport in this area, providing power to electric vehicles along two of the main boulevards in this part of the city. The construction of the substation was driven by the need to expand the traction network due to the growing route network of the tram and trolleybus transport.

As with other rectifier substations operated by Sofia Public Electrical Transport Company JSC, the Hadji Dimitar TRS is built according to a standard single-line principle diagram shown in fig. 1. The substation is supplied by the electrical grid through two incoming high-voltage connections (Incomer 1 and 2) at 10 kV. The medium voltage switchgear (MV SwG) is equipped with SF_6 gas circuit breakers, which, at the time of their implementation, represented a modern technical solution with high reliability. However, in the current context of environmental standards and sustainable development, the use of SF_6 (sulfur hexafluoride) gas is considered incompatible with future requirements for minimal environmental impact [1, 6, 7, 8].

The MV SwG feeds three traction transformers (T-TR), manufactured in Bulgaria by Elprom Energo E.T. Sofia, with identical technical characteristics summarized in Table 1. Each transformer steps down the primary input voltage of 10 kV to a secondary voltage of 520 V. Through a three-phase busbar system, this voltage is supplied to thyristor rectifier modules, which provide the required output voltage of 660 V DC, used for powering the traction network.

Table 1. Technical Parameters of TT Type TMI 2500/10

Technical Parameters of TT – Manufacturer: Elprom Energo E.T. Sofia			
No.	**Parameter**	**High Voltage Side**	**Low Voltage Side**
1	Nominal Voltage, kV	10 +/-5%	0,515
2	Nominal Current, A	126	2455
3	Nominal Power, kVA	2190kVA	
4	Cooling Type	ONAN	
5	Frequency, Hz	50	
6	Weight, kg	7705	
7	Connection Group	Dy-5	
8	Voltage Regulation	10,46	
9	Coil Material	Cu	Cu

As shown in fig. 1, the Thyristor Rectifier Unit (TRU) consists of a total of 13 RU (rectifier units) cabinets, functionally divided into two main groups: common cathode (RU1 – RU6), which is used to supply positive voltage to the catenary network, and common anode (RU11 – RU13), which is connected to the return line feeders of the system. The TRU is manufactured by the Czech company ČKD Praha.

Modules RU1 to RU6 are constructed with NVP-type thyristor blocks, whose anodes are connected to the AC three-phase busbar at a voltage of 515 V AC. Each NVP module contains the positive thyristor half-bridge, forming one half of the three-phase controlled rectifier. The thyristor blocks include three power thyristors (TV978-1600-20), a cooling system (either natural or with forced circulation air), protective and control electronic circuits, as well as visual indicators of their status. The controlled thyristor modules also function as electronic circuit breakers, providing reliable protection and fast disconnection in case of fault conditions. In this type of technical solution, high speed circuit breakers (HSCB) are not used, simplifying the system and reducing the need for complex mechanical protection.

Each NVP-type cabinet supplies two independent sections of the contact network, with a nominal output current of 2x1500 A at 660 V DC.

The RU11 to RU13 cabinets are constructed with NVZ-type thyristor blocks, where the cathodes are connected to the AC three-phase busbar at 515 V AC, forming the second half of the three-phase controlled rectifier. The NVZ modules are the negative thyristor half-bridge, forming the second half of the TRU. The NVZ modules are connected to the return feeders through a current-limiting reactor, whose main function is to limit currents during short-circuit conditions. Each NVZ cabinet is functionally similar to the NVP, containing two thyristor modules rated for a nominal current of 3000 A. Through a system of disconnectors, the thyristor modules are connected to the return bus (- minus).

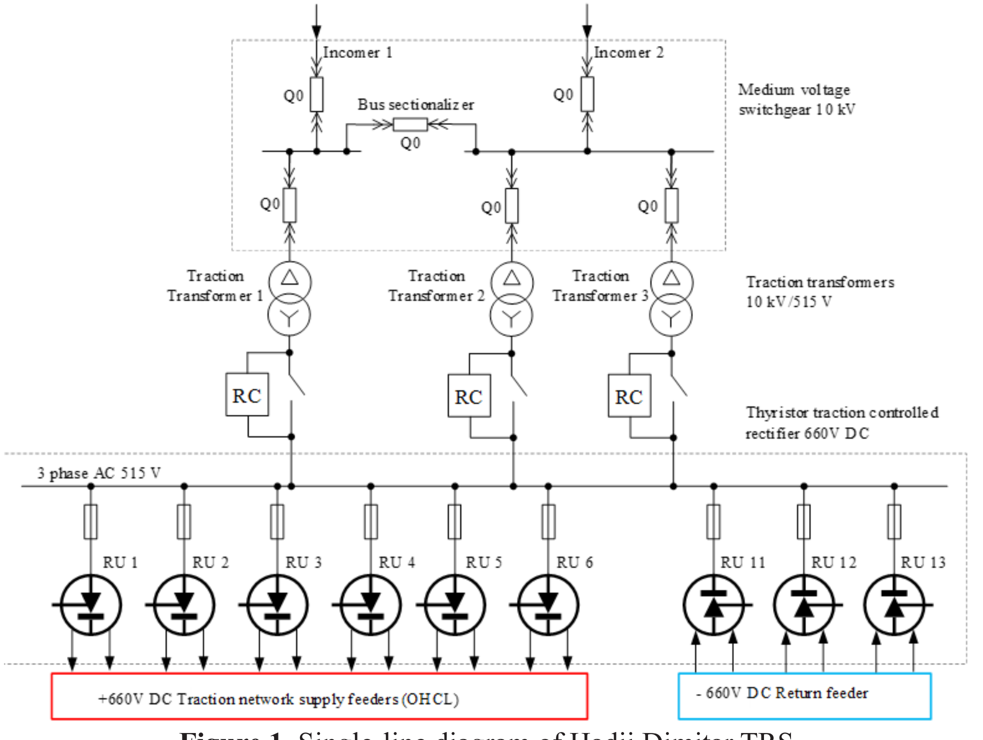
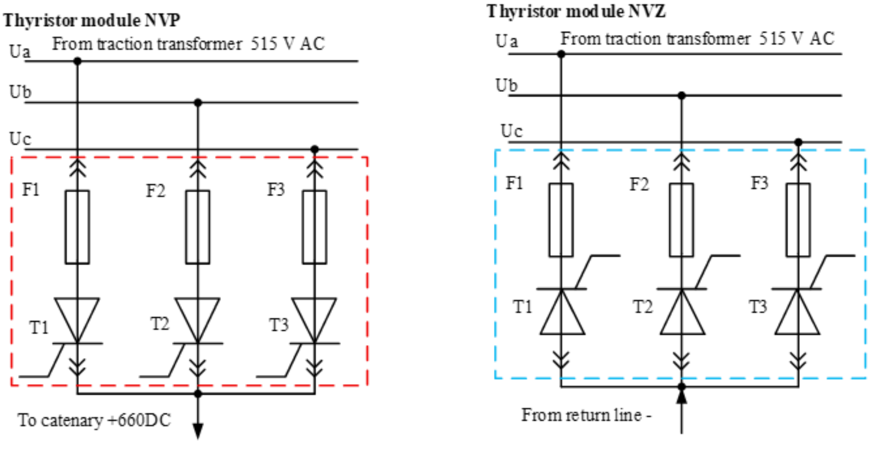


Figure 1. Single-line diagram of Hadji Dimitar TRS

The structure of the power section and the wiring diagram of the NVP/NVZ thyristor blocks are shown in fig. 2.

2.2. Description of the NVP/NVZ Thyristor Block

Figures 2a and 2b show the structural diagrams of two types of thyristor modules used in the TRU of the Hadji Dimitar R.S. – NVP-type module and NVZ-type module. Together, they form a six-pulse thyristor rectifier circuit.



a) **b)**

Figure 2. Single-line diagram of the NVP/NVZ thyristor block

The NVP module (Fig. 2a) is designed for the positive pole of the contact network (+660 V DC). It is powered by 515 V three-phase voltage from the traction transformer. The module consists of three thyristors (T1, T2, T3) with their anodes connected to the alternating current bus, and the cathodes short-circuited to the output. Each thyristor is protected by a fuse (F1, F2, F3).

The NVZ module (Fig. 2b) is intended for connection to the return line (-660 V DC). In this case, the cathodes of the thyristors are connected to the three-phase AC circuit, while the anodes, short-circuited, and are connected to the return feeders. The configuration is a mirror image of the NVP, but functionally equivalent.

Together, the modules provide complete rectification of the three-phase voltage and generate a bipolar DC supply for the electric transport network.

3. Experimental Measurement and Obtained Results

The experimental measurements were conducted on June 15, 2025 (Sunday) at the Hadji Dimitar TRS. The aim of the measurements was to investigate the energy efficiency and operational parameters of the system under real operational conditions.

During the measurement, were operating two traction transformers No. 1 and No. 3. The power supply was provided from Incomer 1 – the intermediate electrical connection with rectifier station Levski. The measurements were taken on a weekend, so it can be assumed that the load on the contact network was lower than usual, especially during the peak hours of a typical weekday. This fact was taken into account when analyzing the obtained values, as the measured parameters did not reach their maximum operational values.

An additional influence on the energy characteristics comes from the current-limiting reactor connected to the -660V bus of the rectifier system.

The measurements were conducted using a Janitza UMG 604 network analyzer [5], which allows simultaneous monitoring and recording of a wide range of electrical parameters. The monitoring was performed simultaneously on all three phases, and the following key parameters were recorded: P, kW – Active power of phase L1, Q, kvar – Reactive power of phase L1, Phase angle of L1, cosϕ of L1, P, kW - Active power for L1÷L3 (summed value for all three phases), Q, kvar - Reactive power for L1÷L3, cosϕ L1÷L3, Power Factor (PF).

The obtained data allows for the evaluation of the energy performance and the efficiency of the thyristor rectifier in specific operating conditions. Below, the report presents summarized results from some of the measurements which characterize the behavior of the rectifier station under limited load conditions.

4. Analysis of Results

Figure 3 presents the graphs of active and reactive power measured at the power input of the Hadji Dimitar Rectifier Substation. These graphs reflect the typical behavior of the power supply when operating trams or trolleybuses. Active power P varies dynamically over time, with clearly visible peaks during acceleration and lower values during coasting or idle periods. The measured minimum momentary value is approximately 13,0 kW (no-load power), and the maximum reaches approximately 544,9 kW.

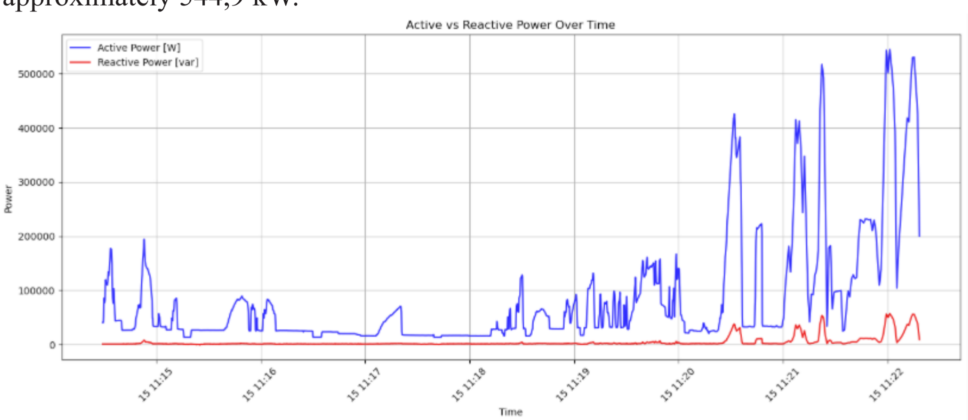


Figure 3. Active and reactive power at the incomer TPSS Hadzhi Dimitar

Reactive power Q also changes according to the load but remains in significantly lower ranges – from about 0,7 kvar (in no-load mode) to 57,0 kvar. Increases in reactive power coincide with moments of higher active consumption, suggesting the presence of inductive elements in the system, such as reactors, transformers, and others.

Figure 4 shows the time dependence of the measured power factor (PF) and cosφ, obtained from the power input of the Hadji Dimitar Rectifier Substation. The measured cosφ remains almost constant in the range from 0,994 to 0,99, which indicates minimal phase displacement between the current II and the voltage U. This corresponds to the expected behavior in an electrical supply with a predominantly active load.

In contrast, the actual PF varies significantly from 0,394 to 0,952, sometimes falling below 0,5. This is a strong indicator of the presence of a nonlinear load, where the current is significantly distorted by higher harmonics. Such values are characteristic of thyristor control, as is the case with the TRU in the TRS Hadji Dimitar.

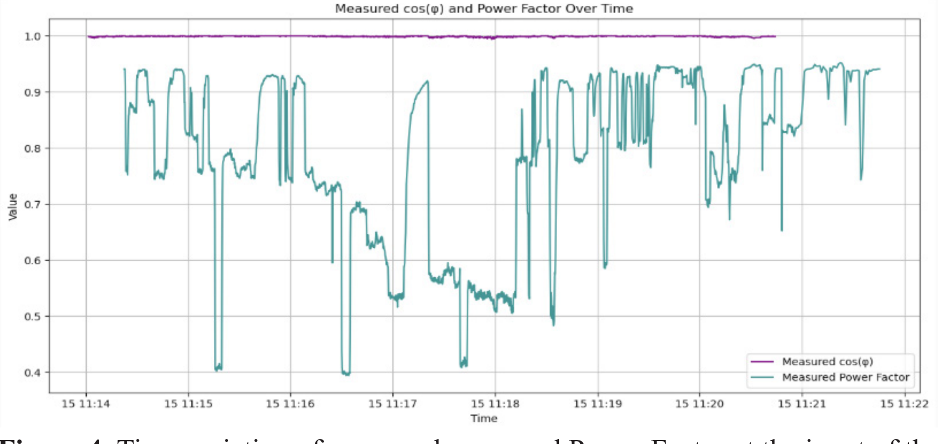


Figure 4. Time variation of measured cos φ and Power Factor at the input of the TPSS Hadzhi Dimitar

The difference between the high cosφ and the low PF clearly indicates that the problem lies not in the phase shift, but in the distortions of the current waveform, which lead to a decrease in the efficiency of the TRU.

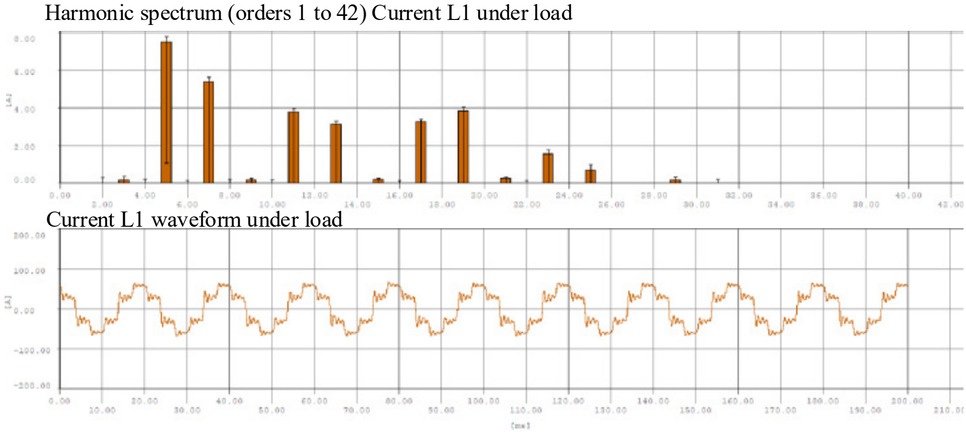


Figure 5. Harmonic Spectrum and Current Waveform of Phase L1 Under Load

Figure 5 shows the spectral composition of the current harmonics and its waveform in phase L1 under load operation of the TRU at the TRS Hadji Dimitar Rectifier Substation. The upper part displays the harmonic spectrum up to the 42nd harmonic order, with the most prominent being the 5th and 7th harmonics, with amplitudes of approximately 7,8 A and 6,0 A, respectively. These are followed by the 11th, 13th, 17th, and 19th harmonics, with amplitudes ranging from 3,5 to 3,9 A. Clearly, higher odd-order harmonics of low order are typical for the operation under load of the TRU. Under this load, the total harmonic distortion (THDi) of the current is approximately 20.5%, indicating the presence of significant higher harmonic components and confirming the characteristic behavior of the TRU under load conditions.

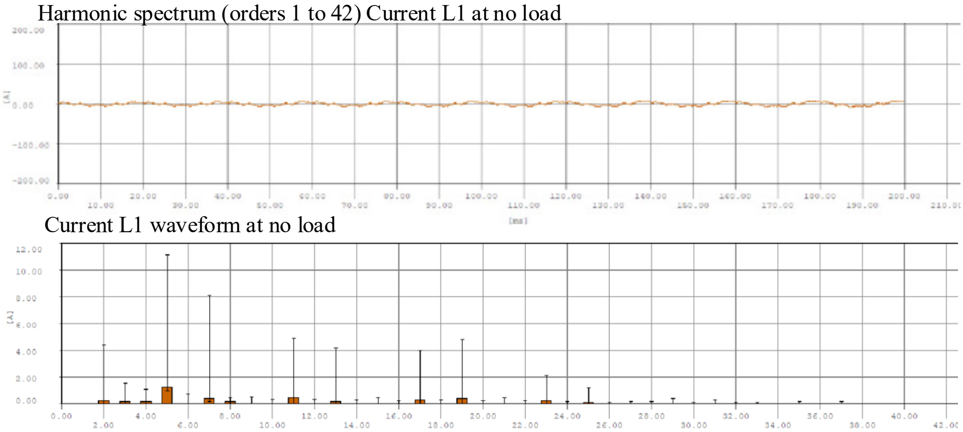


Figure 6. Harmonic Spectrum and Current Waveform of Phase L1 at no Load

When operating without load (figure 6), the most pronounced harmonics are the 3rd (1,2 A), 5th (1,7 A), and 7th (1,1 A), while harmonics from the 9th to the 19th have lower amplitudes – below 1 A. The measured total harmonic distortion (THD) coefficient is approximately 25%, which confirms the presence of significant harmonic pollution even under low load. This indicates that thyristor-driven traction units maintain high levels of harmonic distortion and unfavorable energy parameters even with small loads, in contrast to diode rectifiers, where this phenomenon is much less pronounced.

Conclusion

During the experimental study, real measurements were performed on the main electrical parameters of the TRU installed at the Hadji Dimitar Rectifier Substation. The measured quantities included voltage, current, pulsations, harmonic distortions, cos φ, active and reactive power, as well as the power factor. Based on the obtained data, an analysis was conducted on the current waveforms and the spectral content of the harmonics, both under load operation and in no-load conditions.

The results showed that, despite the good design and the indisputable positive effect of the TRU in stabilizing the output voltage, the system demonstrated unfavorable energy characteristics. The measured total harmonic distortion (THDi) values, reaching up to 20–25%, and the low PF with high cos φ, confirm the presence of significant nonlinear distortions. This leads to additional losses and reduced energy efficiency.

A comparison with diode rectifier devices [3, 4], particularly 12-pulse schemes, shows that the latter offer significantly better energy performance, especially under low and partial loads. In this context, it is both technically and economically justified to consider the phased modernization of existing thyristor-based units with diode or more modern solutions that meet the requirements for energy efficiency and reduced harmonic pollution.

REFERENCES

[1] REGULATION (EU) 2024/573 of the European Parliament and of the Council of 7 February 2024 on fluorinated greenhouse gases, amending Directive (EU) 2019/1937 and repealing Regulation (EU) No 517/2014, Official Journal of the European Union, L 20, 20 Feb. 2024, pp. 1 – 130. Available: https://eur-lex.europa.eu/eli/reg/2024/573/oj.
[2] REGULATION (EU) 2024/1610 of the European Parliament and of the Council of 14 May 2024, amending Regulation (EU) 2019/1242 as regards strengthening the CO_2 emission performance standards for new heavy-duty vehicles, including specific sales targets of 90 % zero-emission new urban buses by 2030 and 100 % by 2035. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401610.
[3] KALUNTA, F., AKINBULIRE, T., OKAFOR, F., Modelling and Analysis of Three Phase Controlled Rectifiers Using Pivotal Function Technique, UNILAG Journal of Medicine, Science and Technology, 2024, vol. 3, no. 1, pp. 82 – 95, ISSN: 2630-693X, https://ujmst.unilag.edu.ng/article/view/130.
[4] LALEV, T., Opportunities for Improving the Energy Performance of the Traction Power Supply System for Surface Urban Electric Transport [Online], Sofia: Author's publication, 2025, 114 p., ISBN 978-619-04-0390-6.
[5] JANITZA ELECTRONICS GmbH, UMG 604 User Manual, Lahnau, Germany. [Online]. Available: https://www.janitza.com/files/download/manuals/current/UMG604/janitza-bhb-umg604ul-en.pdf.
[6] PAVLOV, G., SEKULOV, L., STOITSEV, R., MIHAILOV, L., BODUROV, P., Experimental testing of the recuperative modes of tram motrices PESA 120NaSf along the track segment powered by the 'Emil Markov' and 'Shiroka Laka' feeders of the 'motopista' current-rectifier station in Sofia VI International Scientific Conference – TEIT 2022", Scientific Journal Mechanics Transport Communications, Sofia, Issue 3/2022, art. ID: 2279, pp. X-85 – X-90, ISSN 2367-6620 (online), https://mtc-aj.com/library/2279.pdf.
[7] DODOV, E., PAVLOV, G., Research and analysis of the energy parameters of traction substations of the underground urban electric transport, VI International Scientific Conference – TEIT 2022, Scientific Journal Mechanics Transport Communications, Sofia, Issue 3/2022, art. ID: 2274, pp. X-54 – X-59, ISSN 2367-6620 (online), https://mtc-aj.com/library/2274.pdf.
[8] TARPOV, I., Trends and forecasts in the development of transport concepts for the cities of the future, XXVI international scientific conference TRANSPORT 2023, Scientific Journal Mechanics Transport Communications, Sofia, Issue 3/2023 г., art. ID: 2423, pp. X-1 – X-5, ISSN 2367-6620 (online), https://mtc-aj.com/library/2423.pdf.

✉ **Dr. Todor Lalev, Assoc. Prof.**
ORCID iD: 0009-0006-4214-7942
Department of Electrical Equipment in Railway Transport
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: tlalev@vtu.bg

✉ **Dr. Martina Tomcheva, Assoc. Prof.**
ORCID iD: 0009-0008-8584-4689
Department of Electrical Equipment in Railway Transport
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: mtomcheva@vtu.bg

✉ **Eng. Detelina Pavlova, PhD Student**
ORCID iD: 0009-0008-1660-9167
Department of Electrical Equipment in Railway Transport
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: didaip@vtu.bg

https://doi.org/10.53656/isct-2025.21

AN INTELLIGENT CONTROL APPROACH FOR FREEWAY RAMP METERING

Hassane Abouaïssa, Vasil Dimitrov
Artois University, Béthune, France
Todor Kableshkov University of Transport, Sofia, Bulgaria

Abstract. In the frame of intelligent transportation systems, management of traffic flow in the area surrounding cities and metropolises is an essential aspect of ensuring efficient and smooth transportation. This paper addresses a critical issue in freeway management: the effective regulation of traffic flow to alleviate daily recurrent congestion and ensure the safe circulation of goods and people. It is an effective means of reducing air pollution. The proposed strategy utilizes a two-stage control approach to address freeway ramp metering. The primary objective of ramp metering is to enhance freeway traffic conditions by regulating the inflow from the on-ramp to the highway main lane using signal lights. The first stage relies on differential flatness as a key tool for open-loop control and trajectory planning, offering a strategic approach to ensure operational efficiency. Flat systems are a generalization of linear systems (in the sense that all linear, controllable systems are flat), but the techniques used for controlling flat systems are much different from many of the existing techniques for linear systems. Differential flatness enables system inversion (and linearization) without integrating any differential equation. To handle model inaccuracies, uncertainties, and unexpected traffic demand as well as various disturbances, the second stage introduces model-free control and its corresponding intelligent controllers to close the control loop. This approach enables real-time adjustment of trajectory planning, ensuring that the system reacts effectively to changing traffic conditions. The conducted numerical simulation using real data of the North of France motorway A25, demonstrates the relevance of the proposed approach.

Keywords: Model-free control; differential flatness; freeway ramp metering; intelligent transportation systems

1. Introduction

The worldwide globalization has led to an ever-increasing demand for the transportation of goods and persons. This leading to increased recurrent congestion in the areas surrounding cities and metropolises, particularly during the rush hours.

In this context, intelligent transport systems (ITS) have been developed to ensure, amongst others, efficient management of traffic flows. Because of its crucial impact on the economy, safety and the environment, several measures have been proposed to improve traffic conditions on freeways. Among the most effective measurements, ramp metering has been widely developed and implemented around the word [1]. A ramp meter is a device, typically a traffic light, which uses a signal controller to regulate the flow of traffic entering freeways. Several studies devoted to this important domain of ITS have been proposed (See e.g. [2, 3, 4, 5], and the references therein). These approaches are ranging from classic proportional-integral-derivative (PID) correctors, and more modern control methods (like model predictive control - MPC, or feedback control [6, 7]) to the utilization of ordinary and partial differential equations stemming from mathematical physics, especially from fluid and statistical mechanics (see e.g. [8] and [9, 10]), as well as [11] for a recent survey of traffic flow control. It should be emphasized that ramp metering can be isolated (implemented in the vicinity of each controlled on-ramp) or coordinated (controlling several on-ramps).

Among the many feedback control algorithms which may be found in the huge literature devoted to traffic management, ALINEA (an acronym of the French wording Asservissement LINéaire d'Entrée Autoroutière [12]) is one on the very few closed-loop control syntheses which has been implemented in practice. However, these closed-loop controllers are fixed since they have no means to adapt their gains to changing conditions, and thus suffer from, e.g., uncertainties in traffic parameters and poor tuning. As stated in [1] (see e.g. [13]), one way to solve parameters uncertainties is to estimate them using Kalman filters. However, the assumption underlying the use of Kalman filters is that the parameters follow a known observable model. Moreover, these feedback methods, alongside MPC, often require a mechanism for estimating traffic state in order to respond to real-time traffic behavior.

Recent advances in artificial intelligence (AI) have paved the way for the development of more sophisticated ramp metering approaches. [1, 14] have addressed ramp metering combining MPC and reinforcement learning (RL). Other approaches that have gained substantial attention are based on RL, Deep learning (DL) and Machine learning (ML). However, methods stemming from RL imply control schemes that are tedious to implement [15]. An alternative to these various shortcomings of the algorithms has been proposed in [16]. The fundamental principle underpinning this approach is predicated on the recent advances in model-free control, whereby the mathematical model of the studied system becomes useless [17].

This paper presents work that exploits the power of two control approaches: flatness-based control and model-free control. This results in a two-stage algorithm. The first stage uses differential flatness as a key tool for open-loop control and trajectory planning, providing a strategic approach to ensure operational efficiency. The second stage relies on model-free control in a closed-loop system to address model inaccuracies, uncertainties, and perturbations. The latter, and as demonstrated by [18], is a new tool for ML and is characterized by its ease of implementation and should be substituted in control engineering to ML via Artificial Neural Networks and/or Reinforcement Learning.

This communication is organized as follow. Section 2 provides a brief overview of the macroscopic traffic flow model that is used in this paper. Section 3 is devoted to the main principles of flatness and model-free control, and their application to freeway ramp metering. Computer simulations are conducted in Section 4. Finally, some conclusions and discussions are presented in Section 5.

2. Macroscopic traffic flow model

Macroscopic second-order models are the most widely used models to describe the dynamic of traffic flow along a freeway section. The initial second order model was proposed by Payne [19] and later modified by [20]. The studied model is based on the following conservation equation (as stated by several authors, the only exact equation in traffic flow modeling is the conservation one. All the others equations especially the momentum equation is principally based on empirical considerations and closely related to the specific infrastructure and traffic conditions [21]):

$$(1)\ \frac{\partial}{\partial t}\rho(x,t) + \frac{\partial}{\partial x}q(x,t) = g(x,t),$$

where $\rho(x,t)$ and $q(x,t)$ denote the average aggregate traffic density in units of (veh./km/lane) and the traffic volume in (veh./h), which defines the number of vehicles exiting the segment during the time period dt, respectively; $g(x,t)$ represents the ramp generation term in (veh./h/km).

From a control theory perspective, freeway ramp metering can be viewed as a system controlled by traffic lights near the on-ramp. To derive the control law, consider the given section of freeway depicted in Fig. 1. Let ρ be the traffic density in vehicles per kilometer per lane, q (in veh./h) represents the traffic volume, and v (km/h) is the main speed. According to METANET [22], these three aggregated variable are related by:

$$(2)\ q(t) = \lambda\rho(t)v(t),$$

where λ represents the number of lanes.

The equation of conservation is expressed by the following expression:

$$(3)\ \dot{\rho}(t) = \frac{1}{\lambda L}[q_e(t) - q(t) + q_r(t)]$$

where L is the section length, q_e, q_r are the traffic entering the mainstream section and the on-ramp, respectively.

The dynamic of the mean speed is expressed by:

(4) $\dot{v}_i(t) = \frac{1}{\tau}[V_e(\rho_i) - v_i(t)] + \frac{1}{\lambda L_i} v_i [v_{i-1}(t) - v_i(t)] - \frac{\eta}{\tau L_i} \frac{\rho_{i+1}(t) - \rho_i(t)}{\rho_i(t) + k}$

where τ, η and κ are model parameters. The equilibrium speed V_e is defined by [23]:

(5) $V_e(\rho) = v_f \exp\left[-\frac{1}{a}\left(\frac{\rho(t)}{\rho_{cr}}\right)^a\right]$

where v_f and ρ_c are the free-flow speed and the critical density, respectively; a is a model parameter.

In Eq. (4), the first term represents the relaxation, the second term represents the convexion, and the third term represents the anticipation. It has to be noticed that the above equations represent the basis for the conducted numerical simulations. Equation (5) defines the so-called fundamental diagram shown in Fig. 2.

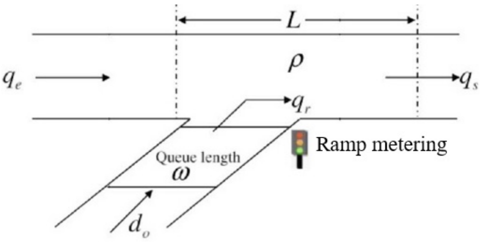


Figure 1. Freeway ramp metering principle

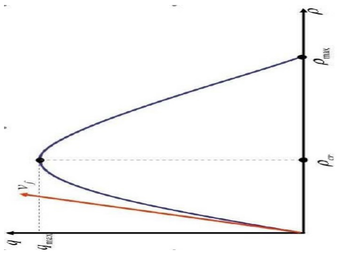


Figure 2. Fundamental diagram

The aim of ramp metering consists on the regulation of the on-ramp flow (using traffic lights) in order to maintain the traffic density at the mainstream segment below a critical one ρ_{cr}.

3. Flatness-based and model-free control

3.1. Differential flatness

Since their introduction more than thirty years ago [24], differentially flat systems have met with great success in control engineering and other fields, such as pure physics [25, 26]. The main principle of this concept can be summarized as follow (interested readers can find more information and theoretical foundation in [27]):

– any system variable z may be expressed as a differential function of the component of the flat output and their derivatives up to some finite order v, i.e.:

(6) $z = f(y_1), \ldots, y_m, \dot{y}_1, \ldots, \dot{y}_m, \ldots, y_1^{(v)}, \ldots, y_m^{(v)}$;

– any component of the flat output may be expressed as a differential function of the system variables;

– the components of the flat output are differentially independent, i.e., they are not related by any differential relation.

Assigning time functions to $y_1,\ldots,y_m$ yields time functions to any system variable without any integration procedure. An open loop control strategy is provided by this, along with allowance for a reference trajectory.

3.2. Model-free control

Faced with the difficulty of obtaining mathematical models for complex systems, Fliess and Join have introduced a new concept in control engineering where the notion of ultra-local model plays an important role. Indeed, consider the SISO (single-input single output) system and replace the poorly known plant and disturbance by the following ultra-local model:

(7) $\dot{y}(t) = F(t) + \alpha u(t)$

where y and u are the output and control variables, respectively; the constant $\alpha \in$ R is often chosen by a practitioner; F, which subsumes all the unknown structure of the studied system including external perturbations, is estimated via algebraic manipulations and the application of the inverse Laplace transform on the interval $[0, T]$ (see e.g. [17] for more explanations). In the time domain, the following data-driven real-time estimator is obtained:

(8) $F_{est}(t) = -\frac{6}{T^3}\int_0^T \big((T-2\sigma)y(t-T+\sigma) + \alpha(T-\sigma)\sigma u(t-T+\sigma)\big)\,dt$

3.3. Ramp metering algorithm

Consider the freeway section of Fig. 1. A natural flat output, derived from Eq. (2) with $\lambda = 1$, is the traffic density ρ. From the principle of the flatness, the open loop control reads:

(9) $u^*(t) = q_r^*(t) = L\dot{\rho}^*(t) + \rho^*(t)v^*(t)$

The trajectory generation is made using a polynomial interpolation. This is accomplished by prescribing the following desired trajectory for the flat output ρ:

(10) $\rho^*(t) = \begin{cases} \rho_0 & \text{for } t < 0 \\ \rho_0 \quad \rho_0 + (\rho_T - \rho_0)\sigma(t,0,T), & \text{for } 0 < 0 < T \\ \rho_T & \text{for } t > T \end{cases}$

where $\sigma(t,0,T)$ is a polynomial function of time, exhibiting a sufficient number of zero derivatives at times, 0 and T, while also satisfying: $\sigma(0,0,T) = 0$ and $\sigma(T,0,T) = 1$.

The control loop is closed thanks to the intelligent controller stemming from model-free control principle

(11) $\Delta u(t) = \frac{1}{\alpha}[-F_{est}(t) + \rho^*(t) + C(e(t))]$

where:

– $e(t) = \rho(t) - \rho^*(t)$ is the tracking error;

– $C(e(t)) = K_P e(t) + K_I \int e(t)\,dt + K_D \frac{de(t)}{dt}$ stands for the classic controller

– K_P, K_I and K_D are tuning gains.

The final control algorithm reads

(12) $u(t) = u^*(t) + \Delta u(t)$

4. Computer simulations

Consider for the computer simulations the section of the North of France freeway A25. The traffic demands at the mainstream and the on-ramp are depicted in Fig. 3. In the no control situation and particularly during the rush hours, around 8 and 10.30 AM, the traffic density increases overhead the critical one. This leads to a formation of congestion and mean speed drop (See e.g. Fig. 4).

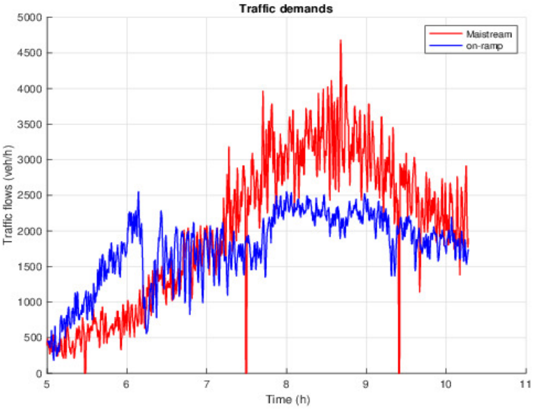


Figure 3. Traffic demands in vehicles per hour

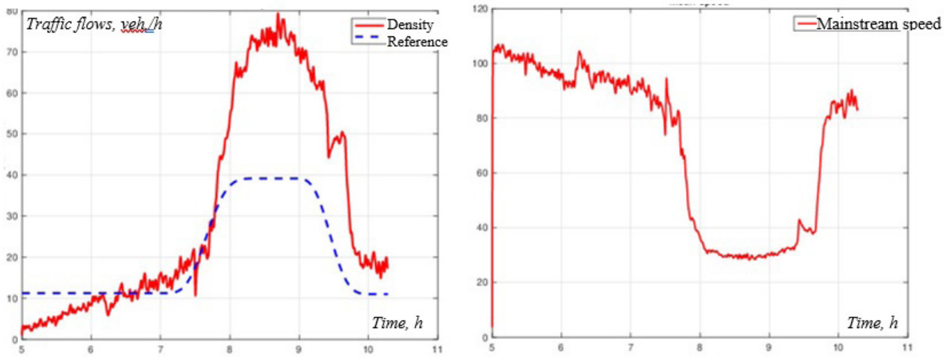


Figure 4. Traffic density and mean speed – no control case

In the no-congestion mode, there is no need for control, the profile of the critical density (the reference) is then low. The controller adapts the density profile to the critical value as soon as traffic becomes dense. Fig. 5 shows significant improvement in traffic conditions.

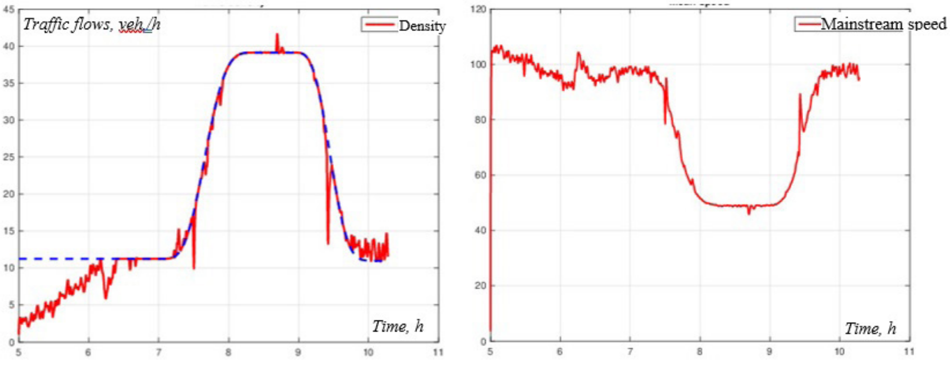


Figure 5. Traffic density and mean speed – control case

Conclusion

The combination of flatness and model-free control seems to be an interesting solution for dynamic ramp metering. Instead of the use of ML techniques which are tedious to implement and require the use of a large amount of data, the application of model-free control and the estimation methods based on algebraic techniques provides an efficient and easy to implement ramp metering approach. This approach enables the dichotomy between the necessity of a model for traffic simulation and the need for a simple, efficient control algorithm to be resolved. However, further investigation must be conducted in order to deal with coordinated ramp metering as well as integrated traffic control where combination with variable speed limits and dynamic traffic routing play an important role.

REFERENCES

[1} CHAVOSHI, K., KOUVELAS, A., Nonlinear Model Predictive Control for Coordinated Traffic Flow Management in Highway Systems. European Control Conference (ECC), St. Petersburg, Russia, 2020, pp. 428 – 433, IEEE, doi: 10.23919/ECC51009.2020.9143962.

[2] ELEFTERIADOU, L., An Introduction to Traffic Flow Theory. 2nd Ed., Springer, 2024, 353 pp., ISBN 978-3-031-54029-5.

[3] GHOSH, S., LEE T.S., Intelligent Transportation Systems: Smart and Green Infrastructure Design. 2nd Ed., CRC Press, 2010, 218 pp., ISBN 9781439835180.

[4] KACHROO, P., ÖZBAY, K., Solution to the User Equilibrium Dynamic Traffic Routing Problem Using Feedback Linearization. Transportation Research Part B: Methodological, Elsevier, vol. 32(5), pp. 343 – 360, ISSN 0191-2615, https://doi.org/10.1016/S0191-2615(97)00031-3.

[5] MAMMAR, S., Systèmes de transport intelligents: Modélisation, information et contrôle. Hermes Science Publications, 2007, 360 pp., ISBN 978-2-746-21519-1.

[6] HEGYI, A., Model Predictive Control for Integrating Traffic Control Measures. TRAIL Thesis Series T2004/2, The Netherlands TRAIL Research School, 2004, 232 pp., ISBN 90-5584-053-X.

[7] KACHROO, P., ÖZBAY, K., Feedback Ramp Metering in Intelligent Transportation Systems. Springer Science+Business Media New York, 2003, 333 pp., ISBN 978-1-4613-4737-8.

[8] KOTSIALOS, A., PAPAGEORGIOU, M., Nonlinear optimal control applied to coordinated ramp metering, IEEE Transactions on Control Systems Technology. 2004, 12/6, pp. 920 – 933, ISSN 1063-6536.

[9] BLANDIN, S., WORK, D., GOATIN, P., PICCOLI, B., BAYEN, A., A general phase transition model for vehicular traffic. SIAM J. Appl. Math., 2011, 71(1), pp. 107 – 127, ISSN 1095-712X, doi:10.2307/41111580.

[10] CHOWDHURY, D., SANTEN, L., SCHADSCHNEIDER, A., Statistical physics of vehicular traffic and some related systems, Physics Rep., Vol. 329/4-6, 2000, pp. 199 – 329, ISSN 0370-1573, https://doi.org/10.1016/S0370-1573(99)00117-9.

[11] SIRI S., PASQUALE, C., SACONE, S., FERRARA A., Freeway traffic control: A survey. Automatica, Vol. 130(5):109655, 2021, ISSN 0005-1098, doi:10.1016/j.automatica.2021.109655.

[12] HAJ-SALEM, H., BLOSSEVILLE, J.-M., DAVÉE, M., PAPAGEORGIOU, M., Alinea: Un outil de régulation d'accès isolé sur autoroute – Étude comparative sur site reel. Rapport INRETS, No 80, Arcueil, 1988, 95 pp., ISBN 2857822421.

[13] SMARAGDIS E., PAPAGEORGIOU, M., KOSMATOPOULOS, E., A flow-maximizing adaptive local ramp metering strategy. Transportation Research Part B: Methodological, Elsevier, vol. 38(3), pp. 251 – 270, ISSN: 0191-2615, https://doi.org/10.1016/S0191-2615(03)00012-2.

[14] HAN, Y., WANG, M., LI, L., RONCOLI, C., GAO, J., LIU, P., A physics-informed reinforcement learning-based strategy for local and coordinated ramp

metering. Transportation Research Part C: Emerging Technologies, Vol. 137/ 103584, 2022, ISSN 0968-090X, https://doi.org/10.1016/j.trc.2022.103584.

[15] DELALEAU, E., JOIN, C., FLIESS, M., Synchronization of Kuramoto oscillators via HEOL, and a discussion on AI. 11th Vienna International Conference on Mathematical Modelling (MATHMOD 2025), Feb 2025, Vienna, Austria. pp. 229 – 234, https://doi.org/10.48550/arXiv.2501.07948.

[16] ABOUAÏSSA, H., FLIESS, M., JOIN, C., On ramp metering: towards a better understanding of ALINEA via model-free control, International Journal of Control, 2017, 90 (5), pp.1018 – 1026, ISSN 1366-5820.

[17] FLIESS, M., JOIN, C., Model-free control. International Journal of Control, Vol. 86/12, pp. 2228 – 2252, 2013, ISSN 1366-5820, https://doi.org/ 10.1080/00207179.2013.810345.

[18] FLIESS M., JOIN C., Machine learning and control engineering: The model-free case, Future Technologies Conference FTC 2020, Vancouver, Canada, Proceedings, Vol. 1, 2020, pp. 258 – 278, ISSN 2194-5357, doi 10.1007/978-3-030-63128-4_20.

[19] PAYNE, H., Models of traffic and control. Simulation Council Proc. Math. Models Public Syst., vol. 1, chap. 6, 1971, pp. 51 – 61, ISSN 0037-5497.

[20] PAPAGEORGIOU, M., BLOSSEVILLE, J.-M., HADJ-SALEM, H., Modelling and real-time control of traffic flow on the southern part of boulevard peripherique in Paris – Part I: Modelling. Transport. Research A, 1990, 24, pp. 345 – 359, ISSN: 1878-3813, DOI: 10.1016/0191-2607(90)90048-B.

[21] SPILIOPOULOU, A., KONTORINAKI, M., PAPAGEORGIOU, M., KOPELIAS, P., Macroscopic traffic flow model validation at congested freeway off-ramp areas. Transportation Research Part C, 41, 2014, pp. 18 – 29, ISSN: 1879-2359.

[22] MESSEMER, A., PAPAGEORGIOU M., METANET: A Macroscopic Simulation Program for Motorway Networks. Traffic Engineering and Control, Vol. 31(8), 1990, pp. 446 – 470, ISSN: 0041-0683.

[23] MAY, A., Traffic Flow Fundamentals. Prentice Hall, Englewood Cliffs, NJ, 1990, 474 pp., ISBN: 0-13-926072-2.

[24] FLIESS, M., LÉVINE, J., MARTIN, P., ROUCHON, P., Sur les systèmes non linéaires différentiellement plats. C. R. Math. Acad. Sci. Paris 315, 1992, no. 5, pp. 619 – 624, ISSN 1778-3577.

[25] FLIESS, M., LÉVINE, J., MARTIN, P., ROUCHON, P., Flatness and defect of non-linear systems: introductory theory and examples. Int. J. Control 61, 1995, no. 6, pp. 1327 – 1361, ISSN 1366-5820.

[26] FLIESS, M., LÉVINE, J., MARTIN, P., ROUCHON, P., Deux applications de la géométrie locale des diffiétés. Ann. Inst. Henri Poincaré, Phys. Théor. 66, 1997, no. 3, pp. 275 – 292, ISSN 0246-0211.

[27] SIRA-RAMÍREZ, H., AGRAWAL, S., Differentially flat systems. Marcel Dekker, New York, 2004, 488 pp., ISBN 0824754700.

✉ **Hassane Abouaïssa, Assoc. Prof., HDR**
ORCID iD: 0000-0002-0017-4110
Laboratoire de Génie Informatique et d'Automatique de l'Artois (LGI2A), Artois University,
UR 3926, F-62400 Béthune, France
E-mail: hassane.abouaissa@univ-artois.fr

✉ **Prof. Dr. Vasil Dimitrov**
ORCID iD: 0000-0001-8206-7521
Department of Electrical Equipment in Railway Transport, Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: vdimitroff@vtu.bg

https://doi.org/10.53656/isct-2025.22

PROJECT-BASED APPROACH TO RISK MANAGEMENT OF TRANSPORT COMPANY INFORMATION SYSTEMS

Mario Bogdanović, Deni Vale
Istrian University of Applied Sciences, Pula, Croatia

Abstract. Information security, manifested as the protection of information and physical data carriers, is a fundamental element in ensuring the business success of transport companies. In order to maintain robust information security in terms of uninterrupted operational use, it is necessary to manage the risks of a given information system. This includes assessing and evaluating its most critical and high-risk components, as well as proposing alternative measures (strategies) for reducing risks (threats) and dealing with specific hazardous scenarios. This paper addresses the issue of achieving appropriate information security through risk management using a project-based methodology. The approach is presented through three defined problems and corresponding hypotheses: (1) Risk identification; (2) Risk quantification with categorization based on severity (high-level/kill risks, medium-level and low-level risks); (3) Determination of the suitability of various alternatives/strategies for managing all types and levels of risk. The results demonstrate that risks can be identified and quantified according to their severity in the information systems of transport companies. Risk mitigation can be approached through one of six (6) risk management strategies (risk avoidance, risk transfer, risk prevention, risk impact reduction, risk ignoring, contingency planning), through their combination/synergy, or by alternative means. The methodological principle of project management in the field of risk management has proven to be a suitable approach for managing risks in the information systems of transport companies and can also be applied to risk management in the information systems of other types of companies.

Keywords: information security; information system; risk management; risk management methodology; risk management strategies; transport companies

1. Introduction

Uncertainty and risk are inherent components of business life, and it is dangerous to ignore or deny their impact, as many things can go “wrong.” Risk management is particularly emphasized in project management, being one of the nine fundamental knowledge areas of the discipline itself [1]. It also draws on creative techniques that employ various modes of thinking to solve complex problems, with one of the key approaches being consideration of potential risks, dangers, damage, and the downsides of ill-considered decisions, which are known as the "black hat" in the Six Thinking Hats technique created by Edward de Bono (for example see Ref. [2]).

Potential threats, negative impacts, and harms arising from uncertainty are described using the terms uncertainty and negative risk. Uncertainty refers to a lack of information, knowledge, or understanding regarding the outcome of an action, decision, or event. Management often suffers from such a lack, which is crucial for effective decision-making. Risk is a measure of the degree of uncertainty and is directly related to the availability of information. The relationship between information and uncertainty is illustrated in Figure 1.

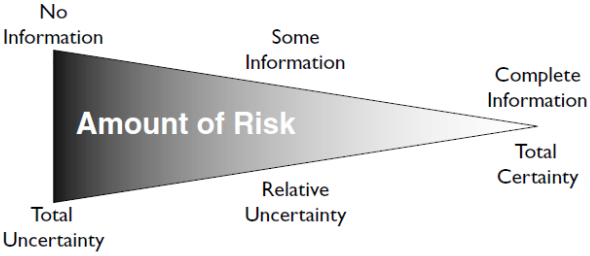


Figure 1. Risk relationship between information and uncertainty, taken from Ref [3]

Effective risk management requires the identification of threats, which are specific events that may lead a business system toward undesirable outcomes such as deadline overruns, cost overruns, poor quality, or failure to maintain business continuity. In digital business, risks arise from the intensive use of information systems and technology, which are essential for supporting and improving business processes and the overall functioning of any enterprise. Digital business risks depend on three variables: the assets of the business entity (both material and financial), threats (incidents, unwanted events), and vulnerabilities (system weaknesses resulting from poorly implemented control mechanisms), hence:

(1) $Risk = f\,(assets, threats, vulnerabilities)$

The manifestation of risk can be either direct or indirect and is typically reflected as a material or financial loss [4]. Material or financial loss occurs in relation to informational assets, which include everything that holds value for an organization, serves as information or contains it, and is used in service support processes based on that information [5]. According to Ref. [6], informational assets encompass all resources that store, transmit, create, use information, or are information themselves.

Information systems of transport companies, regardless of the mode of transportation (automobile, bus, truck, rail, maritime, air, etc.), do not fundamentally differ from the information systems of other companies, as disruptions in these systems lead to similar or even greater negative consequences when security is compromised. In fact, if confusion arises within the information system, business operations and costs are quickly affected. For example, disruptions in timetables or the place/date of dispatch directly impact business efficiency. Therefore, every transport company should have a well-developed risk management scheme for each identified risk, with particular attention to those of higher magnitude.

2. Purpose, Objective, Problems, and Hypotheses of the Research

The purpose of this paper is to present a rational way (methodology) for improving information security in transport companies using a project risk management methodology. The main goal of the paper is to identify the most critical/riskiest parts of the transport company's information system and to propose measures/strategies for reducing certain fatal risks for these risks.

The problems of the work are defined as questions, hypotheses as hypothetical answers to the problems posed:

P1: What are the threatening risks for the information system of transport companies?

H1: It is possible to identify numerous risks of the information system of transport companies.

P2: Which risks of the information system of transport companies can be considered fatal risks (high-risk risks/risks in the high-risk zone), and which are medium and low-risk risks?

H2: It is possible to quantify kill risks (high-level risks) as well as medium and low-level risks.

P3: How to manage (which strategy) each of the quantified information system risks of kill (high), medium and low risk levels?

H3: It is possible to manage the identified and quantified risks using one of six (6) risk management strategies.

3. Research Methodology

For the first problem, the project risk identification methodology from Ref. [3] was used. Using the brainstorming/brainwriting method as well as thinking about potential risks/problems (answering the question: "What could go wrong?"), the risks of the information system were identified. Answers were offered with regard to experience, problems observed in information systems of other types of companies, and hypothetically possible threats.

The formula was used to solve the second problem, i.e. risk quantification:
(2)

$$R = probability\ of\ a\ risk\ event \times\ strength\ of\ the\ risk\ impact\ if\ it\ occurs.$$

The rating scale is 10 points for each element of the product (factor x factor). The results obtained are presented graphically for a clear representation of the risk zone (high, medium, low). The authors of the paper were the evaluators of each risk. The risk zones are presented in Fig. 2.

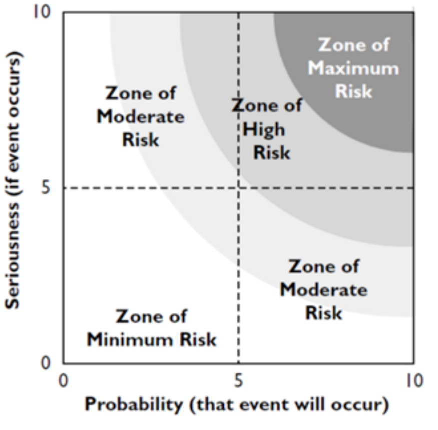


Figure 2. Identification of Risk Zones/Severity Levels, taken from Ref [3]

In the third problem, high (kill risks), medium and low level risks were linked to 6 different strategies. For those risks for which a suitable strategy could not be found, a new strategy, or alternative coping strategy, was proposed.

4. Research Results and Discussion

4.1. Identification and Quantification of Information System Risks in a Transport Company

The list of identified information system risks was obtained using the method of experience and creative production of the authors. A total of *N=44* risks were identified, which are also descriptively explained in an unpublished paper [7]. After identification, all of the listed risks were quantified. This was necessary to distinguish those risks that are high-level – the so-called high-threat/lethal risks – from those of medium and low-threat levels. The identified and quantified information system risks are presented in Table 1.

Table 1. Information system risks, their probability, impact strength and product (danger level)

Types of identified risk	Probability of risk occurrence	The strength of the risk impact if it occurs	Multiplier (danger level). 1=kill (70 – 100); 2=medium (40 – 69); 3=low (<39)
1. Theft of media and documents	4	8	32
2. Unauthorized changes to the system	6	9	54
3. Password security	8	9	72
4. Hacker attacks using specially crafted messages	9	9	81
5. Unauthorized entry into physical space	3	8	24
6. Interception of compromising interference signals	2	7	14
7. Interception of information/ espionage	4	9	36
8. Eavesdropping	5	8	40
9. Error in the application/use of the application	4	7	28
10. Data loss	8	9	72
11. Loss of integrity of sensitive information	1	9	9
12. Mismatched/incompatible IT equipment and resources	5	6	30
13. Electromagnetic interference	1	6	6
14. Flood/moisture/water	2	9	18
15. Fire/excessive temperature/ heat	2	10	20
16. Various natural disasters	1	10	10
17. Main event nearby	1	4	4
18. Failure or interruption of power supply networks/power outage	2	8	16
19. Denial of services (DoS)	5	7	35
20. Failure or interruption of communication networks	6	6	36
21. Coercion, blackmail or corruption	3	7	21
22. Identity theft	5	6	30

Types of identified risk	Probability of risk occurrence	The strength of the risk impact if it occurs	Multiplier (danger level). 1=kill (70 – 100); 2=medium (40 – 69); 3=low (<39)
23. Misuse of personal data	6	7	42
24. Malicious software	7	8	56
25. Denial of harmful actions (sabotage/damage)	4	7	28
26. Unauthorized use or management of devices and systems	6	7	42
27. Abuse of authority/privileges	5	8	40
28. Accidental loss of storage devices/media and documents	4	6	24
29. Sabotage/damage	3	8	24
30. Poor planning or lack of adaptation	4	5	20
31. Disclosure of sensitive information	3	8	24
32. Information or IT products from unreliable sources	3	7	21
33. Hardware or software manipulation	2	8	16
34. Information manipulation	3	7	21
35. Destruction of devices or storage media	2	9	18
36. Device or system failure/ malfunction	4	6	24
37. Pollution, dust, corrosion	3	6	18
38. Lack of financial resources	3	7	21
39. Lack of material and information/IT resources	3	6	18
40. Lack of human resources, i.e. staff / lack of trained personnel	5	7	28
41. Software vulnerability or errors	3	8	24
42. Violation of laws/regulations	2	9	18
43. Adverse climatic influences	2	9	18
44. Social engineering	3	6	18

4.2. Linking Risks with Response Strategies

The identified and quantified threats in the next step are linked to coping strategies, and the main way of implementation is shown in Table 2.

Table 2. Risk response strategies

Strategy	Low-risk threats	Medium-risk threats	High-risk threats
Ignore	Frequently applicable	Caution, only if the costs of protection are too high	Unacceptable, because the consequences could be great
Avoid	Almost not used	Sometimes, if there is an easier alternative	Useful when exposure can be avoided completely
Risk transfer	Rarely, unless already covered by insurance	Common practice (e.g. insurance, outsourcing)	Very common, especially through insurance
Prevention	Useful if it is cheap and simple	Recommended to reduce the likelihood	Mandatory, first line of defense
Impact reduction	If the consequences are known and tolerable	Often used in conjunction with prevention	Required in combination with other measures
Contingency planning	Rarely, unless it is part of a larger plan	Recommended – prepare for the worst-case scenario	Necessary – part of any serious security plan

Avoidance strategies for certain security risks in transport companies are sometimes possible but rarely fully feasible in daily operations. For example, switching to biometric authentication can reduce password-related risks, while spear phishing exposure can be limited through secure communication channels and email filters. However, in cases like data loss, avoidance is not an option due to the essential role of digital data processing in logistics, making prevention and contingency planning crucial. Ignoring risks is generally unjustifiable, and although risk transfer through insurance or outsourcing (e.g., cloud services, Managed Security Service Providers) can reduce the burden of consequences, while reputational and legal damages remain the company responsibility. Risk mitigation strategies for weak passwords include enforcing complex password policies, implementing multi-factor authentication, rotating credentials regularly, and using password managers with breach detection integration. For spear phishing attacks, effective strategies involve targeted staff training, secure email filtering, sender authentication protocols, and link/attachment sandboxing, supported by incident response and containment plans. In the case of data loss, the most robust strategy combines regular immutable backups (3-2-1 rule), access control based on user roles, offline storage, and predefined RTO/RPO parameters to enable rapid system recovery and maintain operational continuity. Here, we have briefly outlined only the strategies for high-risk threats, while a more detailed analysis, including strategies for medium-risk and low-risk threats, is presented in the unpublished study [7].

In this way, all identified and quantified risks are linked to coping strategies, which achieves significant progress in managing risks in information systems of transport companies. Thus, the project approach has proven to be a good tool in managerial risk management of information systems.

Conclusion

Concisely summarizing the obtained results of this work, the following conclusions can be drawn:

P1/H1: Information systems of transport companies are exposed to numerous threats that include technical, organizational, human, legal and natural risks.

P2/H2: Based on a quantitative analysis of the probability of occurrence and the strength of the impact of each of the identified risks, it is possible to clearly distinguish the risks of the information system of transport companies according to their level of danger. The results shown in Table 1. confirm the hypothesis that it is possible to identify kill, medium and low-danger risks. Risks such as hacker attacks with specially crafted messages, password security and loss of business data are classified as fatal because they achieve high product values (above 70), which indicates their extreme harmfulness and high probability of occurrence. Most risks belong to the low-level category, most often due to low probability or limited impact. Such a categorization allows for the adoption of targeted and prioritized security measures, with an emphasis on proactive protection against high-level risks, thereby contributing to the resilience and continuity of information systems within the transport sector.

By linking the identified and quantified information system risks (P3/H3), it can be concluded that effective management of information system risks of fatal, medium and low levels of danger can be achieved by applying one or more of six risk management strategies, possibly their combination/synergy, where the choice of strategy depends on the type and nature of the individual risk. In situations where complete control over the risk is impossible or too expensive, a risk transfer strategy through insurance or contracts with external service providers may be an appropriate solution. Proper application of the strategy requires understanding the specific context and assessing the benefit-cost ratio for each individual risk. This confirms that all identified risks can be managed with a targeted strategy, thereby reducing their harmfulness and increasing the resilience of the information system of transport company.

Acknowledgement

Special thanks go to our colleague Marko Turk, MA Economics, lecturer at the Istrian University of Applied Sciences in Pula, Croatia, for his valuable assistance in compiling the data presented in Table 1. His professional and academic expertise in the field of information systems greatly contributed to the quality of this work, and we sincerely appreciate his support.

REFERENCES

[1] BUBLE, M., Projektni menadžment, Dugopolje (Croatia): Minerva University of Applied Sciences, 2010, 208 pp., ISBN 978-953-56361-0-6.
[2] DOMITRAN, I., Kako koristiti metodu „Šest šešira“? Poslovna učinkovitost d.o.o., 2014, available at: https://www.poslovnaucinkovitost.hr/kolumne/poslovanje/986-kako-koristiti-metodu-sest-sesira (accessed on 14 December 2021).
[3] HEERKENS, G., Project Management, New York, NY (USA): McGraw-Hill, 2002, 250 pp., DOI:10.1036/0071394494. The material in this Book appears as eBook 0-07-139449-4 and print version 0-07-137952-5.
[4] DALZIEL, H., Infosec Management Fundamentals, Amsterdam/Boston: Elsevier Science (Syngress), 2015, 72 pp., ISBN 978-0128041871.
[5] UREMOVIĆ, D., Kako upravljati IT rizicima. InfoTrend, 5 (2009), pp. 42 – 47, ISSN 1330-0393, available at: http://www.infotrend.hr/clanak/2009/6/kako-upravljati-it-rizicima,37,767.html (accessed 14 December 2021).
[6] HADJINA, N., Zaštita i sigurnost informacijskih sustava. Zagreb, Croatia: Faculty of Electrical Engineering and Computing (FER), University of Zagreb, 2009.
[7] BOGDANOVIĆ, M., VALE, D., A Detailed Project-Based Analysis of Information System Risks in the Transport Sector, unpublished manuscript, Croatia.

✉ **Prof. Dr. Mario Bogdanović, scientific adviser**
ORCID iD: 0000-0003-3049-5342
Department for Social Sciences
University of Applied Sciences in Pula
9D, Preradovićeva, Pula, Croatia
E-mail: mbogdanov@iv.hr

✉ **Deni Vale, mag. phys., lecturer**
ORCID iD: 0000-0002-3138-7581
Department for Technical Sciences
University of Applied Sciences in Pula
9D, Preradovićeva, Pula, Croatia
E-mail: dvale@iv.hr

https://doi.org/10.53656/isct-2025.23

Telecommunications and Signaling Equipment, Automation Systems in Transport

RESEARCH AND ANALYSIS OF THE OPERATIONAL RESULTS AFTER THE IMPLEMENTATION OF A TRAIN CONTROL SYSTEM WITH GRADE OF AUTOMATION GOA3 ON THE METRO LINE 3 IN SOFIA

Svetoslav Tomov, Emilia Dimitrova
Todor Kableshkov University of Transport, Sofia, Bulgaria

Abstract. With its three metro lines in the city of Sofia, the operator Metropolitan EAD takes the central share in the portfolio of the capital's urban transport. The implementation of a CBTC system for signaling and train operation on the third metro line in Sofia is undoubtedly the most high-tech transport project ever happened at the territory of the Republic of Bulgaria. The challenges for the implementation of this project are enormous both for the employer – Metropolitan EAD, and for the civil works contractors and in particular for the contractor of the signaling systems. This unique technological leap for the capital's urban transport had a strong reflection on the national level, placing Bulgaria second in the world and first in Europe with the implementation of the specific software package for the constituent set of subsystems. The main goal is a safe, fast and comfortable transport service for the residents and guests of the capital. This process is accompanied by several new requirements, a need to reconsider the current conventional model and introduce changes in the daily operation, maintenance and management of the urban railway infrastructure. This report provides an analysis of the results in terms of operation, changes in the operator profile and passenger behavior several years after the introduction of urban rail transport with a GoA3 automation level train control system in Sofia. Looking ahead to potential future upgrades to the line and the corresponding additional benefits that can be achieved through them are provided.

Keywords: urban rail transport; passengers' safety; CBTC; GoA3; signaling system; automatic train control

1. Introduction

The automated metro line three in Sofia represents a significant step towards sustainable and intelligent urban transport. It is expected to improve the quality of life, reduce the environmental footprint and increase the efficiency of the transport system in the capital. The key indicators for the quality of the transport service provided are the safety and comfort of passengers, its reliability and speed of movement. Of course, the introduction of technology of such a level for the first time in the country, as the pilot project in terms of the set of systems for automatic train control and line management, has its challenges. The main signaling system implemented on metro line three in Sofia is based on IEEE 1474.1 Standard for Communications-Based Train Control (CBTC) Performance and Functional Requirements [1]. In matter of automatic train operation (ATO) the grade of automation is GoA3. This level of automation is considered as a Driverless Train Operation (DTO) where the personnel still accompany the train but can move away from the cabin [2]. However, at the request of the operator – Metropolitan – the trains have been delivered with separate cabins with a driver control panel and a chair. Therefore, the doors and the departure of the train from the platform are controlled manually from the cabin. With the section of the first eight stations being put into operation in 2020, and then in 2021 four more stations added, the third metro line in Sofia is still considered young. The operation of the line itself and its control systems have not been studied yet. The purpose of the report is to explore the trend of behavior of the line control systems and how automation changes the behavior of passengers and operating personnel. This document provides an analysis of the operational results of the interaction of the set of different signaling subsystems implemented on the line, and the new order that automation leads the transport service to.

2. Material and Methods

This article is based on observations made during the actual passenger operation of the line and examines the factors that affect the transport service. The transport service directly depends on the availability and reliability of the main signalling subsystem: Automatic Train Control (ATC) including its Radio/Data Communication System and Interlocking system (IXL), Automatic Train Supervision (ATS), Power Supply/ Distribution (PDC), Platform Screen Doors (PSD), Public Information System (PIS), and the actions of operational and maintenance staff and the behavior of passengers. The scope covered by this report is shown in *Figure 1 Areas of observation covered by the report.* The trains are only considered with their onboard signalling equipment. All of these could affect the presence of the service and its smoothness.

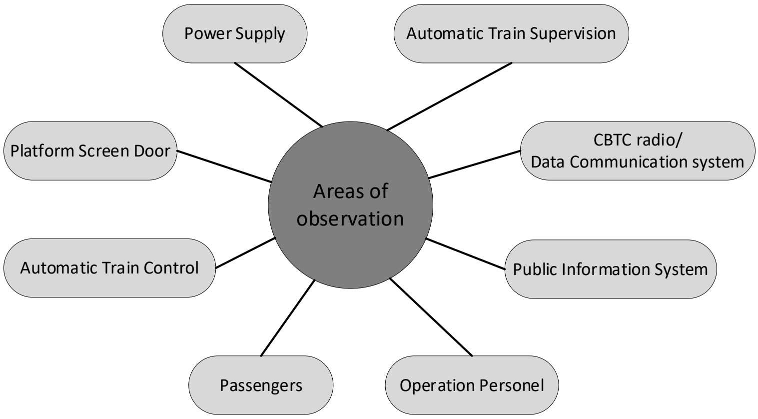
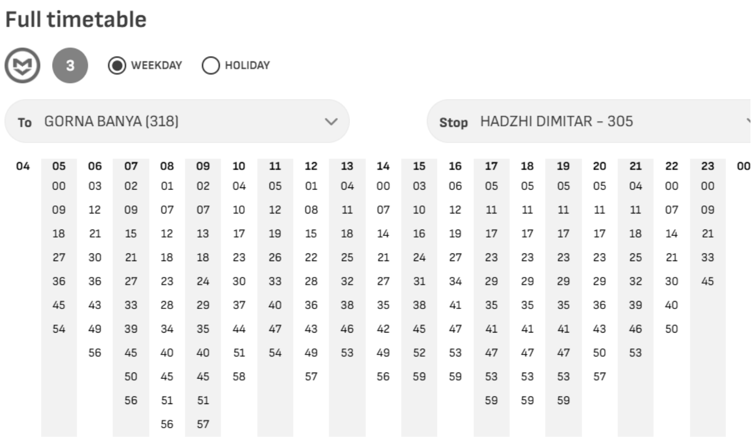


Figure 1. Areas of observation covered by the report

The headway (the time interval between two successive trains moving along the same track in the same direction) offered by the Automatic Train Control (ATC) system is just 90s [3]. The public operation on metro line three in Sofia is available from 05:00h to 23:45h. The rest of the time the line is used for planned or corrective maintenance, training and other non-passengers' related activities. Due to the current needs of the operator, the headway used is up to 5 min, see *Figure 2 Weekday timetable of M3 line in Sofia, July 2025* [4]. However, to follow the timetable strictly there are additional conditions to be considered.



04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00
	00	03	02	01	02	04	05	01	04	00	03	06	05	05	05	05	04	00	00	
	09	12	09	07	07	10	12	08	11	07	10	12	11	11	11	11	11	07	09	
	18	21	15	12	13	17	19	15	18	14	16	19	17	17	17	17	18	14	21	
	27	30	21	18	18	23	26	22	25	21	24	27	23	23	23	23	25	21	33	
	36	36	27	23	24	30	33	28	32	27	31	34	29	29	29	29	32	30	45	
	45	43	33	28	29	37	40	36	38	35	38	41	35	35	35	36	39	40		
	54	49	39	34	35	44	47	43	46	42	45	47	41	41	41	43	46	50		
		56	45	40	40	51	54	49	53	49	52	53	47	47	47	50	53			
			50	45	45	58		57		56	59	59	53	53	53	57				
			56	51	51								59	59	59					
				56	57															

Figure 2. Weekday timetable of M3 line in Sofia, July 2025 [4]

Failures of single components reduce the redundancy but normally should not have an operational effect. However sporadic inconsistencies between different software applications running in parallel could lead to looping and blocking of the entire system. The information about the subsystem's hardware availability has been obtained via the implemented dedicated Service and Diagnostic (S&D) tool. It is a software application that indicates the hardware alarms for both wayside and onboard signalling equipment and proposes corrective actions to the maintenance personel. Fluctuations in system reliability could come from the simultaneous operation of different but interconnected software applications. The observations about these cases are based on cases registered during the warranty and post-warranty support.

Despite the automation of the line, the human factor, as far as operational and maintenance personnel, as well as the passengers themselves, can cause a malfunction in service. Any time an operator (e.g., train delocalization) or technician intervention is required (e.g., platform screen door issue), there is a risk of misjudged actions that could exacerbate the problem and further delay according to the timetable. Automated operations must guaranty passengers' safety [5]. As a safety strategy measure there is a half-height platform screen door (PSD) structure at all stations of metro line three. Their main function is to guarantee the passengers' safety as they have a role of a physical barrier to the track side. In addition, the PSD structure is also supposed to reduce the piston effect when trains enter the station and hold the wave in case of flood from the track side [6, 7]. Trains' precise positioning according to the openings of the PSD structure is important for maximum comfort during the passenger's exchange [8]. Passengers on Metro Line 3 in Sofia got used to recognize the opening parts of the PSD structure and to group waiting around them. This optimizes the waiting time of the trains so that the schedule is not delayed.

On Metro Line 3, trains must stop at a fixed point on the platform (a stopping window) with an accuracy of +/- 30 cm [1]. *Figupe 3 Stopping accuracy of train B015 on 04 July 20025 at Krasno selo metro station, platform 1,* provides a shot of precise train stopping. However, this might not be valid all the time. This is an important moment for the operation, because if a train stops outside the stopping window the train doors will open automatically but platform doors will not. They will wait until the operator opens them manually via the system. If a passenger opens a platform door from the emergency handle at the train side and then leaves this door open, this will lead to a need an operational personel to go to the platform to close it. Until then the train will stay standstill, this will cause delays for this train and very likely for the next coming. The observations here were collected by monitoring of such similar situations during drives and operation. Of course, here the error can become cumulative. For example, when doors on a given platform need to be manually opened, the on-duty station operator opens the doors on the opposite platform. In this case, all trains in the area of the station and its entrances and exits sections will be automatically stopped by the CBTC system, as there is a possibility that a passenger will fall onto the railway tracks. This is already considered as a safety-related case. Such an operation disruption decreases the comfort of the transport service.

Figure 3. Stopping accuracy of train B015 on 4 July 2025 at Krasno selo metro station, platform 1

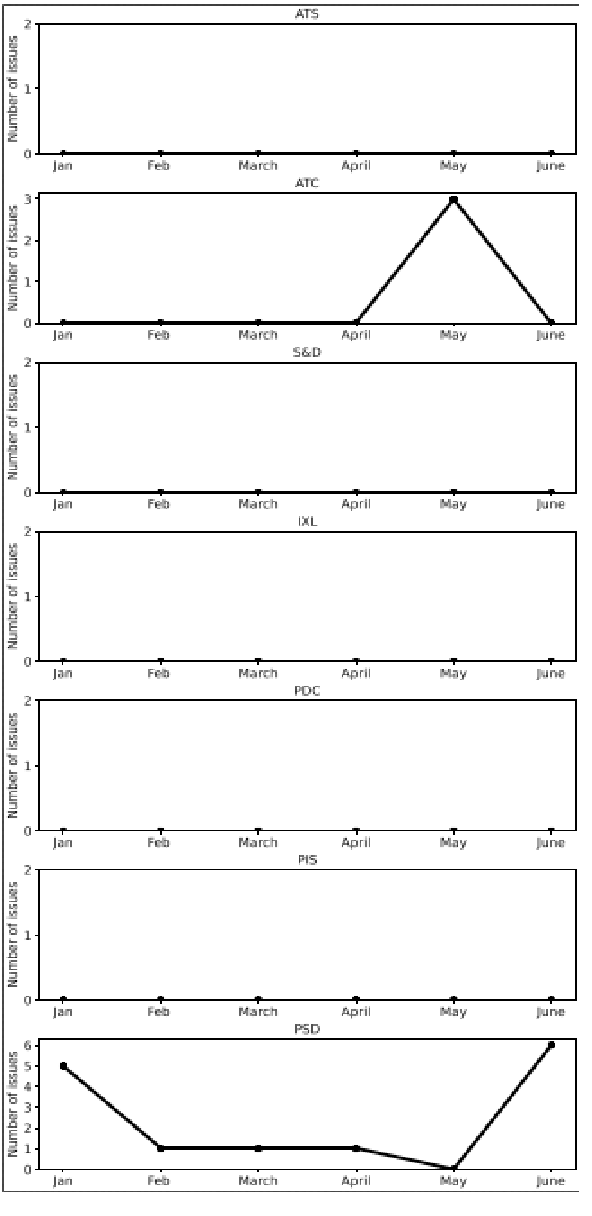
Also, in case of obstacles to the platform doors, the time needed for verification by the system extends the train's dwell time as per the following formula:

(1) $T = 2 + N8 *$,

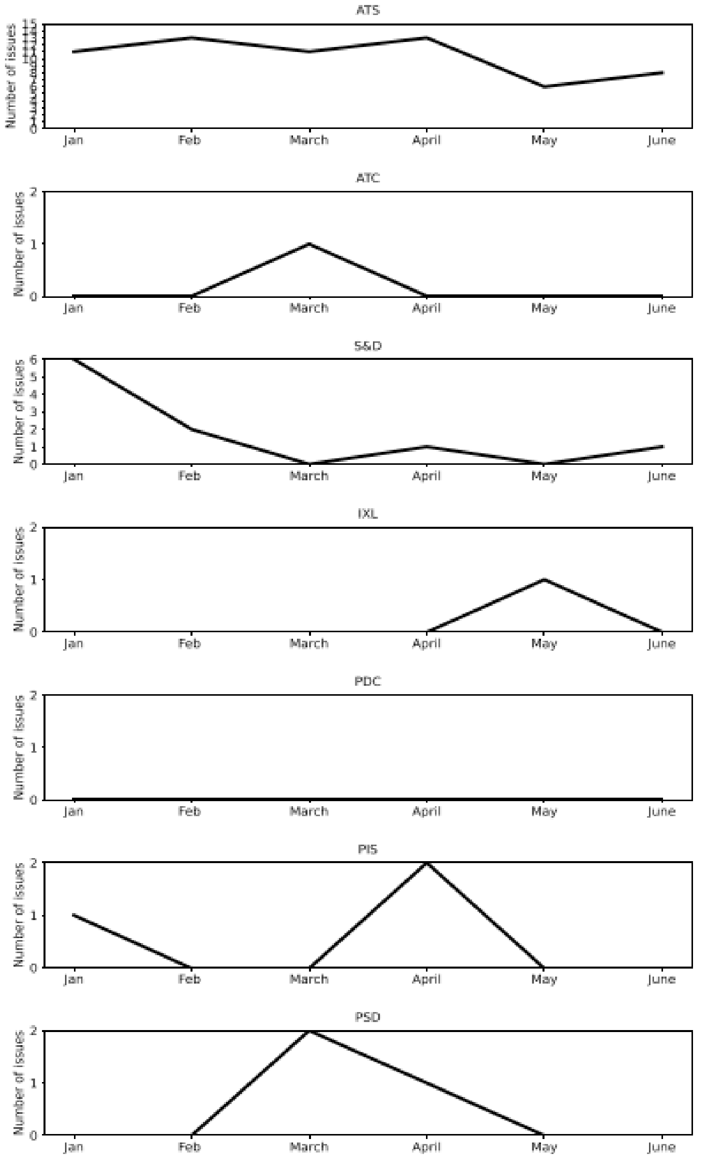
where "T [s]" is the train's dwell time in seconds and "N" is the number (1, 2, 3) of a platform door reopening and closing in case of obstacle. After the third unsuccessful attempt to close the door, the system considers there is something and remains open waiting for an operational personel to come and check it on the spot.

3. Results

Usually at the very opening of the public operation certain hardware may fail more often and some software applications may need more adjustments. The third metro line in Sofia was no exception to this. But more important is how the system behaves after this period has passed. The statistics for hardware availability and software reliability for the first six months of line operation for 2025 are shown in *Figure 4 Overview of signaling systems behavior for M3 line in Sofia, January – July 2025.* High level of availability the system is achieved by using the full redundancy for both the track side and the on-board equipment. As communication with trains is crucial for the line operation, the radio of the CBTC signalling system is required to achieve availability higher than 99,98%. In fact, the operation of the line has never stopped. This means that key indicator has been achieved at the time of the report. It should also be noted that the planned 100% passenger safety and no incidents have been also met. Given these circumstances, the focus of the observation of transport service on the third metro line falls on the ability to adhere to the time schedule and passenger comfort.



a) Hardware availability



b) Software reliability

Figure 4. Overview of signaling systems behavior for M3 line in Sofia, January – July 2025

The stopping accuracy depends on the tracks conditions, atmospheric influences, the train's passenger load, the tolerances in the sensors and systems of a given train, the selected train operation mode. Stopping accuracy measured on line three in Sofia is shown in Table 1. Where the values are with a "+" sign, it means that the train has passed the stopping point, i.e. the place where the train doors would be centered relative to the location of the platform doors. Values with a "-" sign mean that the train has stopped before the stopping point. In any case all measured values are within the norms of the stopping window (+/- 30 cm) and allow normal passenger exchange.

Table 1. Stopping accuracy of train B013 measured on 29th of July 2021

Metro Station	Platform 1	Platform 2
05	Not in use for the moment	+18 cm
06	+16 cm	+14 cm
08	+13 cm	+16 cm
09	+13 cm	-1 cm
10	+16 cm	+3 cm
11	+10 cm	+6 cm
12	+21 cm	+7 cm
14	+17 cm	+7 cm
15	+9 cm	+5 cm
16	+7 cm	+3 cm
17	+11 cm	-1 cm
18	+8 cm	+13 cm

The observed characteristic operational advantages and disadvantages of the signaling system solution for the third line of the Sofia metro are listed in Table 2.

Table 2. Line M3 in Sofia signaling systems operational observations

Advantages	Disadvantages
The hardware is extremely stable. Also, exchanging a replaceable unit when needed takes a little time and can be done without interfering with operation.	Moving parts of the PSD needs regular maintenance. If this is ignored, it will lead to a disrupted transport service and delays.
The implemented pilot software package allows easy monitoring and control of the metro line and supports the work of maintenance personel.	Software is an alive system. The updates close some issues but often bring new ones.

The operational staff is provided with a highly functional system with multiple capabilities, including event recording and playback.	High qualified personel is required. A mistake when solving a small problem can dramatically increase the severity of the situation.
Passengers are guaranteed 100% safe and comfortable trips	Passengers' abnormal behavior could cause disrupted transport service and delays.

Conclusion

In general, the signaling system of metro line three in Sofia is a high-tech product that meets the requirements for safety, preventing the occurrence of critical situations. The availability and reliability of the transport service seems to be limited to a much smaller scale of hardware failures and replacement. More often, adjustment or restart, tightening is required. More often, restarts, adjustments or simply fixation tightening are required. An important factor is the quality and conditions of the initial installation - dustiness of the room, working without air conditioning, vibrations, etc. A problem with the software of the system is often reflected in a delay in the time schedule. For example, train delocalization requires the train to move at a limited speed of 20 km/h to the next section, and sometimes even longer. Given the automation of the line, the dependence of the quality of the transport service on the behavior of the operator's staff and on the passengers themselves should be low. However, the fact is that they have their importance and in practice this cannot be ignored. It requires the presence of highly qualified personnel with constant maintenance of knowledge and the implementation of refreshing training. The goal from the passengers' perspective should be to achieve higher awareness and culture of behavior.

With the future line extension/ branching it is expected that on a certain section on the line there will be an alternate movement of trains for two different destinations. This will lead to decreasing the headway into half [9].

As a next step in the development of the line three control system, future upgrades with already available systems for automatic analysis of aggregate operational information can be envisaged, which provides advice on optimizing line management and maintenance and offers the operation personel preventive actions at an earlier stage, before corrective actions are required.

REFERENCES

[1] METROPOLITAN EAD, Volume 5: Technical specifications. Employer's requirements, Section 5.8: CBTC - Communication Based Train Control, Sofia, 2019, pp. 19, 26.

[2] SCHNIEDER, L., Communications-Based Train Control (CBTC) Components – Functions – Operations, Germany, 2019, pp. 28 – 29, ISBN 978-3-96245-200-1 (PMC Media House: Bingen am Rhein).

[3] SIEMENS MOBILITY, CBTC – Trainguard MT Enhancing passenger experience through migration and integration of digital services, Germany, 2021, p.3.
[4] URBAN MOBILITY CENTER, Virtual timetable/Sofia, 2025, https://www.sofiatraffic.bg/en/public-transport.
[5] SCHNIEDER, L., Protection of Critical Transport Infrastructure – A holistic approach, Germany, 2019, pp.33-36, ISBN: 978-3-96245-168-4 (PMC Media House GmbH: Bingen am Rhein).
[6] METROPOLITAN EAD, Volume 4: Technical Specifications. Part 4.12: System for automatic platform screen doors (SAPSD), Sofia, 2021, pp.5, 8.
[7] DIMITROV, V., Study of the possibilities for optimisation of positioning of electric vehicles, In: Proceedings of the 12th Electrical Engineering Faculty Conference (BulEF), Varna, Bulgaria: IEEE, 2020, p. 3, ISSN 2603-5359.
[8] PAVLOV, N., DIMITROV, V., Influence of the braking on the comfort during positioning of a metro train. In: Proceedings of the 2019 11th Electrical Engineering Faculty Conference (BulEF). Varna, Bulgaria: IEEE, 2019, pp. 1 – 4. ISSN 2603-5359.
[9] METROPOLITAN EAD, Volume 4: Technical Specifications. Employer's Requirements, Part 4.8: Communication Based Train Control (CBTC), Sofia, 2021, p. 19.

✉ **Svetoslav Tomov, Assist. Prof.**
ORCID iD: 0009-0002-6109-0137
Department of Telecommunications and Safety Equipment and Systems
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: tomov_svetoslav@abv.bg

✉ **Prof. Dr. Emilia Dimitrova**
ORCID iD: 0000-0001-6813-0563
Department of Telecommunications and Safety Equipment and Systems
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: edimitrova@vtu.bg

https://doi.org/10.53656/isct-2025.24

RELIABILITY ASSESSMENT OF RAILWAY SIGNALING SYSTEMS USING FAULT TREE ANALYSIS

Vasil Ivanov, Emiliya Dimitrova
Todor Kableshkov University of Transport, Sofia, Bulgaria

Abstract. The report presents the application of the fault tree analysis method in evaluating the reliability of equipment included in railway signaling systems. To calculate the reliability indicators, the algebra of random events is applied, where failures of individual components are treated as such events. The study aims to identify the relationships between failures of individual components and the overall (final) system failure through logical modeling of their interdependencies. Fault tree analysis provides a symbolic representation of the conditions that may lead to a system failure, using logical operators to define the relationships between individual component failures and the final failure. This approach helps identify weak points, assess redundancy, and verify compliance with the specified safety and reliability requirements. As an example of logical modeling using fault tree analysis, an automatic level crossing system (ALCS) was examined. The analysis aims to present the relationship between failures in the barrier mechanism, the control circuitry, and the track sensors, and how they may lead to a common system-level failure – in this case, the crossing remaining open when a train is approaching. Based on the constructed fault tree and predefined failure rate values, the main reliability indicators related to the final system failure were calculated: probability of failure-free operation, probability of failure, failure rate, downtime factor, availability factor. The methodology enables assessment of the current state of these indicators and supports planning, technical maintenance, and informed decision-making for the optimization of railway infrastructure systems. The method is applicable both to existing systems and at the design stage.

Keywords: fault tree analysis; reliability; railway transport; failure rate; downtime factor; availability factor

1. Introduction

The interconnection of technical components (e.g., railway transport facilities) in terms of reliability can be realized in three ways: series, parallel, or mixed. In series connection, the failure of a single component leads to the failure of the entire system. In parallel connection, all components must fail for the system to fail. Mixed connection is obtained by combining serial and parallel connection. Reliability calculation can be carried out by formulating a mathematical model of the system or by means of a functional diagram. The second method is applicable only when the elements in the system are connected in series. After compiling a model of the technical system, the stage of performing analysis follows. During the analysis, mathematical expressions must be defined, allowing for the direct substitution of experimental data in order to calculate the system's reliability indicators [1].

The purpose of constructing fault trees is to trace which specific failures (of components or subsystems) can lead to failure of the entire system. The construction of these trees is achieved by logically connecting the individual events – the failures of the basic elements – using the operators "AND" and "OR".

It is characteristic of the logical "OR" operator that it is used to represent series-connected elements. When the failure rate of the initial elements is constant, the resulting failure rate represents their sum. When using the probabilities of failure-free operation $P(t)$, the resulting probability is the product of the probabilities of the initial elements.

The logical operator "AND" is used to represent parallel connected elements. When the failure rate of the initial elements is constant, the output failure rate is a function of time. When using the probabilities of failure, the resulting probability is the product of the probabilities of the initial elements. Therefore, when calculating the logical "AND" operator, unlike the logical "OR" operator, the probability of failure $Q(t)$ is used instead of the probability of failure-free operation.

When combining the above-mentioned operators, a third logical operator is obtained – the quorum function. It is characteristic that a precisely defined number of basic elements M, out of all N elements, must fail in order for the entire system to fail. When $M = 1$, the quorum function is equivalent to the "OR" operator, and when $M = N$, the operator corresponds to "AND" [2].

When calculating the downtime factor, the following formulas are used:

– For the logical "OR" operator:

(1) $K_D = 1 - \prod_{i=1}^{n} \frac{\mu_i}{\lambda_i + \mu_i}$

– For the logical "AND" operator:

(2) $K_D = \prod_{i=1}^{n} \left(1 - \frac{\mu_i}{\lambda_i + \mu_i}\right)$,

where K_D is the downtime factor, n is the number of elements, included in the input of the logical operator, λ_i is the failure rate of the i-th basic element, μ_i is the repair rate of the i-th basic element.

When calculating the availability factor, the following formula is used:

(3) $K_A = 1 - K_D$

When *calculating* failure rates, the following formulas are used for the logical operators:

– For the logical “OR” operator:

(4) $\lambda_p(t) = \lambda_1(t) + \lambda_2(t) + \ldots + \lambda_n(t)$, if $\lambda_1(t) = \lambda_2(t) = \ldots = \lambda_n(t)$ = const, then

(5) $\lambda_p(t) = \lambda_p = n.\lambda$,

where λ_p denotes the failure rate of the system under consideration as a whole (the top-level failure), expressed as a function of the failure rates of the individual components in the fault tree structure. The index *p* comes from the English word “primary” [3].

– For the logical “AND” operator:

(6) $\lambda_p(t) = \{\sum_{i=1}^{n} \lambda_i (Z_i - 1)\}\{\prod_{j=1}^{n} Z_j - 1\}^{-1}$, where:

(7) $Z_j = \frac{1}{1-e^{-\lambda_j t}}$

In reliability theory, it is accepted that for every technical element or system, the probability of failure-free operation *P*(*t*) and the probability of failure *Q*(*t*) always sum to 1. This is because the two events are considered complementary, mutually exclusive, and collectively exhaustive of all possibilities [4]:

(8) $P(t) + Q(t) = 1$

2. Calculation of Reliability Indicators using the Fault Tree

2.1. Description of the Fault Tree Diagram

An automatic level crossing system (ALCS) is considered, with the top event being the failure of the crossing to close when a train approaches (Fig. 1). Three events are presented, and a fault tree has been constructed. Normally, when the level crossing is open, the road signal relay (RSR) is activated. The condition for closing the level crossing is the occupation of any of the pre-crossing sections, as a result of which the RSR releases and activates the light and sound signals. The barrier mechanism is controlled by a closing relay (CR) – it is normally activated and the barriers are raised. When the conditions for closing the level crossing are met, it releases and triggers the lowering of the barrier gates. When the conditions for opening the level crossing are met, the CR is activated again, and the barriers are raised. The RSR is also reactivated when all conditions for opening the level crossing are met, i.e., when the train has cleared the crossing and the barriers are raised. To register the actual passage of the train through the crossing and to activate the RSR and CR, a normalizing relay (NR) is also included in the circuit. Two auxiliary direction relays (even and odd) are also used to determine the direction of the train's movement [5].

The train's entry into a given pre-crossing section is detected by a track sensor. The presence of a train can be detected using track circuits or point sensors. In the first case, two track circuits cover the two pre-crossing sections, with one acting as the entry circuit for the train movement and the other as the exit circuit. To register the train's entry, the track relay of the entry circuit is released, which closes the level crossing. To open it, a third track circuit is required, which controls the level crossing itself. Failure to activate any of the track circuits leads to untimely opening of the crossing when accidental occupation and release of sections of the track occur [6, 7].

In the second case (the level crossing has point sensors), four sensors are required. Two of them correspond to the train entering a given pre-crossing section, while the other two monitor the actual passing of the train.

If any of the sensors fail to activate, it leads to untimely opening of the level crossing.

In the constructed fault tree, the first event is assumed to be the failure of the track sensor to activate.

The second event is a dangerous failure in the control circuitry – manifested as a violation of the condition for closing the level crossing when the pre-crossing section is occupied. The RSR remains activated [8, 9].

A logical "AND" operator is used for the two events, as both failures must occur in order for a train to enter without the level crossing being closed.

The third event leading to the final failure is included through the logical "OR" operator – failure in the barrier mechanism. The cause of the failure may be a malfunction in the CR, which fails to send the closing command, or a failure in the barrier's drive mechanism [10, 11].

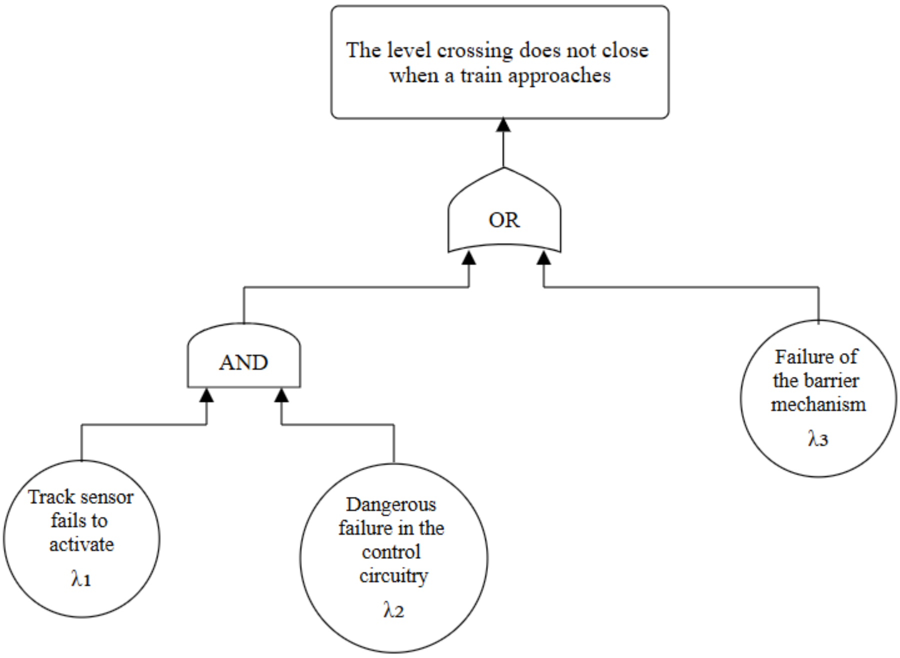


Figure 1. Fault tree representing the conditions leading to a specific failure in the ALCS

2.2. Algorithm for calculating reliability indicators

The basic elements are described by an exponential distribution, using the formula:

(9) $P(t) = e^{-\lambda t} e^{-\lambda t}$

The following indicators need to be determined:

i. The probability of failure-free operation of the technical object for 1300 h – $P(1300)$ = ?, using formula (9) (The probability of failure-free operation of the top event from the fault tree shown in Fig. 1 is to be determined). The initially given values for the basic elements are as follows:

$$\lambda_1 = 13.10^{-5}\ \frac{1}{h}, \lambda_2 = 17.10^{-5}\ \frac{1}{h}, \lambda_3 = 21.10^{-5}\ \frac{1}{h};$$

ii. The failure rate for an operating time of $t = 1000$ h, if the initially given values for the basic elements are as follows:

$\lambda_1 = 11.10^{-5}\ \frac{1}{h}$, $\lambda_2 = 15.10^{-5}\ \frac{1}{h}$, $\lambda_3 = 27.10^{-5}\ \frac{1}{h}$ using the formulas (4), (5), (6) and (7);

iii. The downtime and availability factors of the top event, assuming the repair rate is μ = 0,05 $\frac{1}{h}$ and the initially given values for the basic elements are:

$\lambda_1 = 47.10^{-4}\ \frac{1}{h}$, $\lambda_2 = 53.10^{-4}\ \frac{1}{h}$, $\lambda_3 = 47.10^{-5}\ \frac{1}{h}$ using the formulas (1), (2) and (3);

2.3. Calculation Results

Performed calculations for the three groups of indicators:

i. The calculation is performed in parts, as there are two blocks. In the first block, the logical operator is “AND” – “The resulting probability of failure $Q(t)$ is equal to the product of the failure probabilities of the basic elements”. For this reason, the probabilities of failure and failure-free operation of the basic events are first determined:

– Probability of failure-free operation of the first basic event $P_1(t)$:

$$P_1(t) = e^{-\lambda_1 t} = e^{-13.10^{-5}.1300} = e^{(-0,169)} \approx\approx 0,84$$

– Probability of failure of the first basic event $Q_1(t)$:

(10) $Q_1(t) = 1 - P_1(t) \approx\approx 1 - 0,84 \approx\approx 0,16$

(The formula is a direct application of the algebra of events, as the events “failure” and “failure-free operation” are considered complementary events) [12].

– Probability of failure-free operation of the second basic event $P_2(t)$:

$$P_2(t) = e^{-\lambda_2 t} = e^{-17.10^{-5}.1300} = e^{(-0,221)}e^{(-0,221)} \approx 0,80$$

– Probability of failure of the second basic event $Q_2(t)$:

$$Q_2(t) = 1 - P_2(t) \approx 1 - 0,80 \approx 0,20$$

After calculating the probabilities of failure and failure-free operation of the basic elements, the probabilities for the two branches of the tree and the top event are also determined.

First block (branch):

Since the logical operator is “AND”, the resulting probability of failure is the product of the failure probabilities of the initial elements.

(11) $Q_{block\,1} = Q_1(t)\,.\,Q_2(t) \approx$ 0,16 . 0,2 $\approx$ 0,032

$P_{block\,1} = 1 - Q_{block\,1} \approx$ 1 – 0,032 $\approx$ 0,968

Second block (branch):

It contains only one event, so the probability of failure-free operation is calculated directly:

$P_{block\,2} = e^{-\lambda_3 t} = e^{-21.10^{-5}.1300} = e^{(-0{,}273)} \approx$ 0,76

Finally, the probability of failure-free operation of the top event is determined:

With the logical "OR" operator, the values of the failure-free operation probabilities of the two blocks are multiplied [13]:

(12) $P_{\text{Top event}} = P_{block\,1}\,.\,P_{block\,2} \approx$ 0,968 . 0,76 $\approx$ 0,736

ii. The failure rate is to be determined for an operating time of t = 1000 h

For Block 1, with a logical "AND" operator, formulas (6) and (7) are used:

From formula (6) it follows: $\lambda_{block\,1}(t) = \lambda_{block\,1}(1000) = \frac{\lambda_1(z_1-1)+\lambda_2(z_2-1)}{z_1 z_2 - 1}$

The two basic events have the following values: $\lambda_1 = 11.10^{-5}\,\frac{1}{h}$

and $\lambda_2 = 15.10^{-5}\,\frac{1}{h}$

From formula (7) it follows: $z_1 = \frac{1}{1-e^{-\lambda_1 t}} = \frac{1}{1-e^{-11.10^{-5}.1000}}$

$= \frac{1}{1-e^{-0{,}11}} = \frac{1}{1-0{,}896} \approx$ 9,62

$z_2 = \frac{1}{1-e^{-\lambda_2 t}} = \frac{1}{1-e^{-15.10^{-5}.1000}} = \frac{1}{1-e^{-0{,}15}} = \frac{1}{1-0{,}86} \approx$ 7,14

$z_1.\,z_2$ = 9,62 . 7,14 = 68,68

$\lambda_1.(z_1 - 1) = 11.\,10^{-5}$. 8,62 $\approx$ 9,48. 10^{-4}

$\lambda_2.(z_2 - 1) = 15.\,10^{-5}$. 6,14 $\approx$ 9,21. 10^{-4}

$\lambda_1.(z_1 - 1) + \lambda_2.(z_2 - 1)$ = 9,48. 10^{-4} + 9,21. 10^{-4} = 18,69. 10^{-4}

The final result is:

$$\lambda_{block\,1}(t) = \lambda_{block\,1}(1000) = \frac{\lambda_1(z_1-1)+\lambda_2(z_2-1)}{z_1.z_2-1} = \frac{18{,}69.10^{-4}}{68{,}68-1} = 2{,}76\,.10^{-5}\ \frac{1}{h}$$

The failure rate of Block 2 is given by the problem statement, as it contains only one basic element:

$$\lambda_{block\,2}(t) = \lambda_3 = 27.10^{-5}\ \frac{1}{h}$$

The failure rate of the top event is formed as the sum of the failure rates of the two input blocks, since there is a logical "OR" operator. As indicated in formula (4), the individual failure rates will be added:

$$\lambda_{\text{Top event}}(t) = \lambda_{block\,1}(t) + \lambda_{block\,2}(t) = 2{,}76.\,10^{-5} + 27.\,10^{-5} =$$

$$29{,}76.\,10^{-5}$$

iii. The two blocks are considered again, and formulas (1), (2), and (3) are used: For Block 1, the operator is logical "AND":

$$K_D^{block\,1} = \left(1 - \frac{\mu}{\mu+\lambda_1}\right).\left(1 - \frac{\mu}{\mu+\lambda_2}\right) =$$

$$\left(1 - \frac{0{,}05}{0{,}05+0{,}0047}\right).\left(1 - \frac{0{,}05}{0{,}05+0{,}0053}\right) \approx 0{,}00823$$

In Block 2, there is only one basic event, and its downtime factor is:

$$K_D^{block\,2} = \frac{\lambda_3}{\mu+\lambda_3} = \frac{0{,}00047}{0{,}05+0{,}00047} \approx 0{,}009312$$

The downtime factor of the top event is:

$$(13)\ K_D^{\text{Top event}} = 1 - \left(1 - K_D^{block\,1}\right).\left(1 - K_D^{block\,2}\right) = 1-(1-0{,}00823)\,.\,(1-$$

$$0{,}009312) = 0{,}01746$$

The availability factor of the top event, according to formula (3), is:

$$K_A^{\text{Top event}} = 1 - K_D^{\text{Top event}} = 0{,}9826$$

Analysis of the Results

The calculated probabilities of failure-free operation $P(t)$ and of failure $Q(t)$ for the basic events in the fault tree reveal a distinction between elements with varying levels of reliability within the system. The probability of failure-free operation for the first basic event is approximately 0,84, for the second it is 0,80, and for the third it is lower – 0,76. This draws attention to the third component, which in this case is the barrier mechanism.

The probability of failure for the first block, obtained through the logical "AND" operator, is relatively low – 0,032, which results in high reliability of this branch, $P(t) \approx 0{,}968$. The final probability of failure-free operation of the ALCS (the top event), obtained through the logical "OR" operator, is approximately 0,736. This means that, given the specified parameters, there is approximately a 74% probability that the ALCS will operate without failure up to the defined moment. Accordingly, there is about a 26% risk of failure.

The obtained values confirm the effectiveness of the fault tree method for identifying weak links and highlight the need for targeted improvements in elements with high failure rates.

For Block 1 (logical "AND" operator), the value of the downtime factor is approximately 0,00823, and for Block 2 it is approximately 0,009312.

For the final event, a downtime factor of approximately 0,01746 was obtained, and the corresponding availability factor is approximately 0,9826. This means that, at the given moment, the system is capable of performing its function about 98% of the time.

Conclusion

The fault tree method is distinguished by its high clarity, the ability to clearly present the logical dependencies between system components, and the comparative ease of calculations. Compared to Markov models, which require a complex transition matrix and often assume equal probability of states, fault trees allow for flexible description of different scenarios through the use of logical operators. The method is particularly suitable for the design phase, as well as for the analysis of existing systems with partial or aggregated data [13].

REFERENCES

[1] GINDEV, E., Introduction to the Theory and Practice of Reliability. Part 1. Fundamentals of Applied Reliability. Sofia: Prof. Marin Drinov Publishing House, 2000, pp. 74 – 245, ISBN: 954-430-608-0 (4.1).

[2] THEEG, G., VLASENKO, S., et al. Railway Signalling and Interlocking. 3rd edition. Hamburg: PMC Media House GmbH, 2020. ISBN 978-3962451779.

[3] PACHL, J., Railway Signalling Principles. Braunschweig: Technische Universität Braunschweig, 2020.

[4] MORGAN, C., A System of Automatic Train Control. Norderstedt: Creative Media Partners, LLC, 2022. ISBN 978-1016245411.

[5] IVANOV, E., DIMITROVA, E., Automatic Control of Traffic. First edition., Sofia: Todor Kableshkov University of Transport, 2012, pp. 156-166, ISBN: 978-954-12-0223-4.
[6] HUANG, C., EL HAMI, A.; RADI, B., Overview of Structural Reliability Analysis Methods – Part I: Local Reliability Methods, ISTE OpenScience, ISTE Ltd. London, UK, 2016.
[7] ROSS, SH., Introduction to Probability Models, In Introduction to Probability Models, 13th Ed., Elsevier Inc. 2024, ISBN: 9780443187612.
[8] BASTIDAS-ARTEAGA, E.; SOUBRA, A.-H., Reliability Analysis Methods. In Stochastic Analysis and Invers Modelling, Grenoble, 2014, pp. 53 – 77. ISBN 978-2-9542517-5-2.
[9] TERJE, A., Introduction to reliability analysis, In Risk Analysis, John Wiley & Sons Ltd., 2015, ISBN 978-1-119-05779-6.
[10] KUMAR, V., SINGH, L., TRIPATHI, A. K., Reliability analysis of safety-critical and control systems: a state-of-the-art review, IET Software, 2018, vol. 12, no. 1, pp. 1 – 18. ISSN 1751-8806. DOI: 10.1049/iet-sen.2017.0053.
[11] EFANOV, D., OSADCHY, G., KHÓROSHEV, V., New Stage in Safety Traffic Control Technologies Development: Digital Railroad Crossing, 2019 International Russian Automation Conference (RusAutoCon), Sochi, Russian Federation, IEEE, 2019, pp. 1 – 6. ISBN 978-1-7281-2731-1. DOI: 10.1109/rusautocon.2019.8867700.
[12] YUN, J., TANG, W., Automatic Threshold Adjustment for Predictive Level Crossing Sampling Data Converters, IEEE 67th International Midwest Symposium on Circuits and Systems (MWSCAS), Springfield, MA, USA, pp. 1033-1035, ISBN 979-8-3503-8718-6, doi: 10.1109/MWSCAS60917.2024.10658886, 2024.
[13] EFANOV, D., KHÓROSHEV, V., OSADCHY, G., Principles of Safety Signalling and Traffic Control Systems Synthesis on Railways, 2023 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Sochi, Russian Federation, pp. 634 – 638, ISBN 978-1-6654-7595-2, DOI: 10.1109/icieam57311.2023.10139292t.

✉ **Dr. Vasil Ivanov, Assist. Prof.**
Department of Telecommunications and Signalling
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: vivanov@vtu.bg

✉ **Prof. Dr. Emiliya Dimitrova**
ORCID iD: 0000-0001-6813-0563
Department of Telecommunications and Signalling
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: edimitrova@vtu.bg

https://doi.org/10.53656/isct-2025.25

Telecommunications and Signaling Equipment, Automation Systems in Transport

EVALUATION OF THE CHARACTERISTICS AND COVERAGE AREA OF THE LORAWAN NETWORK

Emil Iontchev[1], Rashko Vladimirov[1], Rosen Miletiev[2], Tsvetan Petrov[1]

1) Todor Kableshkov University of Transport, Sofia, Bulgaria

2) Technical University, Sofia, Bulgaria

Abstract. The article presents a project of LoRaWAN network designed to implement a "smart irrigation system". The network architecture includes two remote end modules connected to a standard gateway. The freely available The Things Network (TTN) is used as a network server. The data are received and commands are sent through the application server implemented on the Raspberry Pi platform. The received data are stored in the specialized InfluxDB database. Their visualization in graphical form is performed with Grafana. A Semtech calculator is used to theoretically determine some of the important parameters of the network, such as its performance and energy consumption. The limits for the size of the payload are calculated in such way to meet the TTN server's requirement for the network usage with a selected expansion factor and bandwidth. The aim of the article is to investigate possible ways to improve individual network parameters to obtain maximum range with a minimal energy consumption from the batteries of the end devices. The experiments are conducted to validate the values of individual parameters obtained from the calculations. The influence of the antenna position of the end devices on the coverage area is also studied. For this purpose, the TTN Map platform is used, on which the coverage areas with the corresponding signal values are visualized on a real map (with real coordinates). Two different scenarios of antenna placement are implemented. A mobile phone is used to obtain the geographic coordinates, with the TTN Mapper application pre-installed on it. The application is connected to the network server using the MQTT protocol, from where it receives the data packets which are sent by the end devices. The data with added coordinates is sent back to the TTN server, from where it is automatically transmitted to the TTN Map server using the Webhooks method via the HTTP protocol. The calculated parametric values are compared with the obtained experimental values, and conclusions are drawn about possible options for the network expansion.

Keywords: LoRaWAN network; internet of things; wireless communications; microcontroller; TTN mapper

1. Introduction

The integration of the capabilities of radio frequency identification (RFID) for automatic recognition and tracking of objects with the capacity to transmit relevant information via the Internet infrastructure, with a view to improving logistics processes, leads to the formulation of the concept of the Internet of Things (IoT), introduced by Kevin Ashton [1]. The development of information and sensor technologies, cloud computing, wireless networks, and the transition to the IPv6 protocol made it possible to talk about "intelligent systems". In them, individual "smart things" are connected in a network and can communicate with external systems via the Internet. Each node of the network is able to perceive or influence the environment, process the received data and wirelessly transmit or receive information. A set of rules for an exchanging information between nodes in a reliable and secure manner has been created, which has led to various communication protocols. The main characteristics by which these protocols are compared and evaluated are: range of action; data transmission speed; energy consumption; security; interoperability; ability to easily expand the network; network topology; price. According to the data transmission range, the protocols can be divided into two groups, short-range and long-range wireless communication protocols. The first group of protocols is implemented mainly based on the 802.15.4 standard [2], the most famous are: ZigBee; 6LoWPAN; Thread; HART; Wi-SUN; etc. The main thing about them is that they self-organize in a network, the communication between the nodes is each with each other and the data transmission rates are low. The second group is united under the name low-power long-range (LPWAN) protocols [3]. The protocols of this group can be divided into two subgroups depending on whether the frequency range they use requires a license or not. The first group is the protocols using existing cellular networks, but adapted for IoT purposes, the most famous are: LTE-M; NB-IoT; LTE Cat 1; and LTE Cat 1bis. The second group of protocols operates in frequency ranges that do not require a license, while providing a range comparable to cellular. The most famous representatives are: Sigfox; LoRaWAN; MIoTy; Weightless; etc. With LoRaWAN, networks with a large range of operation can be implemented at a relatively high data rate and low power consumption. These characteristics are due to the patented signal spread spectrum technology from the company Cycleo, which was later purchased by the company Semtech, which produces receivers and transmitters using the technology. Assessment of the quality and range of point-to-point connections using only the physical layer of LoRaWAN are considered in [4, 5, 6]. The application of the LoRaWAN protocol in various areas and assessment of its characteristics is made in [7, 8, 9, 10, 11]. A summary of the methods for research and assessment of the characteristics of LoRa and LoRaWAN in various IoT applications is made in [12].

The article discusses the main parameters and dependencies of the LoRaWAN protocol, which must be taken into account when designing a specific network. On this basis, a specific architecture has been selected in order to implement an "intelligent irrigation system". The calculations of some of the parameters, which are necessary to meet the requirements of the public network server and network setup, have been performed. The experiments have been conducted to evaluate the coverage range of the gateway using different polarizations of the end device antenna.

2. Basic dependencies for the physical layer of the LoRaWAN protocol

In Europe, LPWAN uses the unlicensed frequency band 865 to 870 MHz. It is divided into six sub-bands, which are subject to different combinations of the two main restrictions applied to this type of network – the transmission power and the channel occupation time (duty cycle) within one hour. The first channel is in the range 865 – 868 MHz with a maximum allowable transmission power of 25 mW and a duty cycle of 1%, which means that a node can transmit 36s within one hour, requirements set by ETSI EN300.220-2 V3.2.1. Additionally, to increase noise immunity and the ability to use the same channel simultaneously by several users, spread spectrum technology is applied. For this purpose, the symbols are formed from a different number of linear frequency modulation (LFM) pulses, in the range of the selected channel bandwidth. The number of bits of the useful signal that are transmitted with one symbol is called the spreading factor (SF). Depending on its value, the number of different symbols, which are used with the corresponding SF, is determined by the equation:

(1) $different\ symbols = 2^{SF}$

The duration of transmission of one symbol is determined by the equation: [13]

(2) $T_s = \frac{2^{SF}}{B}$

Where B is the signal bandwidth. As many bits of the useful signal are transmitted, the duration of the symbol transmission is longer, and the symbol bitrate is lower. Since one symbol has SF bits, the transmission rate of one bit of information is determined by [13]:

(3) $R_b = SF\frac{B}{2^{SF}}$

The addition of check bits to the information bits for the detection and recover transmission error purposes results in a further reduction in transmission speed, determined by the coefficient - coding rate (CR). It is determined by the ratio of the number of information bits to the total number of bits sent.

The transmission time of one LoRa frame (Time on Air) depends on the selected frame configuration. To facilitate the calculation of this time, there is a calculator on the Semtech website [14]. With a given LoRa frame configuration, the number of bytes with useful information, the code rate, SF, and bandwidth, the required frame transmission time is calculated, as well as the channel frequency parameters. These are the received signal strength (RSSI), the signal-to-noise ratio (SNR), and the channel budget (Link budget). To determine the last parameter, it is necessary to know the sensitivity of the receiver used and the maximum transmitter power when we receive its highest value. For the 868MHz frequency range, the maximum permissible transmitter power is 14dBm (25mW). With a receiver sensitivity of -130dBm, the channel budget is 144 dBm.

3. Architecture of the researched LoRaWAN network

A hybrid LoRaWAN network is used to implement a "smart irrigation system". The end devices, the gateway and the application server are managed by the service user, and the network server is managed by the service provider The Things Network [15]. The basic architecture of the system includes two MKR WAN 1310 communication modules. Data reception and processing is performed by a 32-bit Arm® Cortex®-M0 SAMD21 processor. The LoRa connection with the gateway is carried out with the CMWX1ZZABZ module from Murata. The LoRa transceiver used in the module is SX1276. The ECC508 module is used to ensure protection of the transmitted information. One end module is located directly next to the area with plants, a soil moisture sensor and a humidity and air temperature sensor are connected to it. The second is located in close proximity to the water tank, to which the drip irrigation system is connected. This module includes a sensor for measuring the water level in the tank, a solenoid valve for starting and stopping water to the drip irrigation system, and a relay for turning on the pump for filling the tank with water. It also includes an alarm for notification in case of problems with the tank. The block diagram and algorithm of the system are shown in Figure 1.

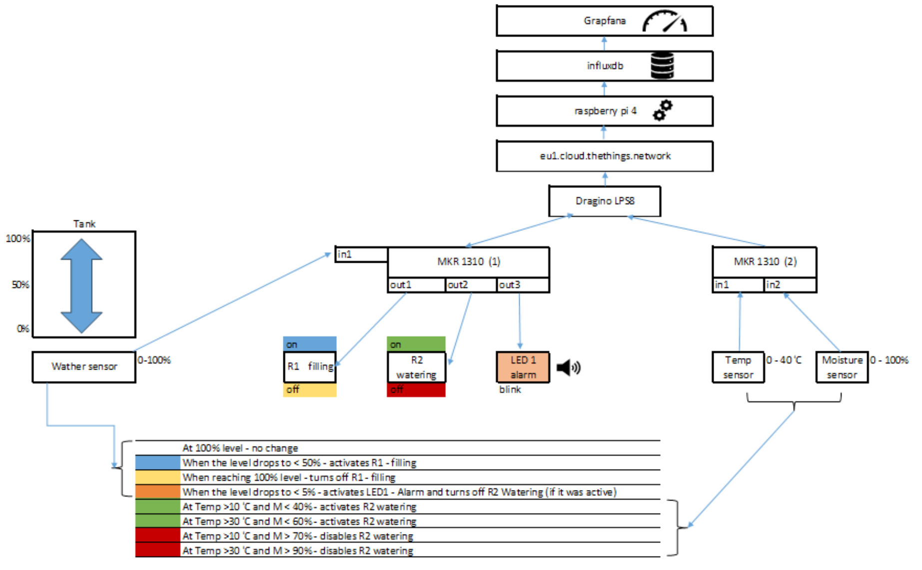


Figure 1. Block diagram of an "intelligent irrigation system"

The gateway is the Dragino LPS8, which provides 10 programmable parallel demodulation channels based on the SX1308. Its location depends on the distances between the two communication modules and is the subject of research. The connection to the Internet depends on its location and can be stationary or mobile. The application server is implemented on the basis of the Raspberry Pi single-board computer. The collected data is stored in a specialized InfluxDB database, and Grafana is used for their graphical presentation. The stored data can be used for statistics for past periods and, if necessary, adjustments can be made to the system settings.

4. Survey the gateway coverage area

The public The Things Network, which is freely accessible, was used as a network server. In order to use it, the restrictions introduced by the administrators were taken into account, namely the network connection time was limited to 30 seconds per day (24 hours) per node, and the messages from the server to the end device were limited to 10 messages per day (24 hours) [15]. This was done by using the calculator [14] to calculate the required Time on Air depending on the parameters that can be set with the software for managing the end devices and the gateway. The values of the individual parameters are shown in Figure 2.

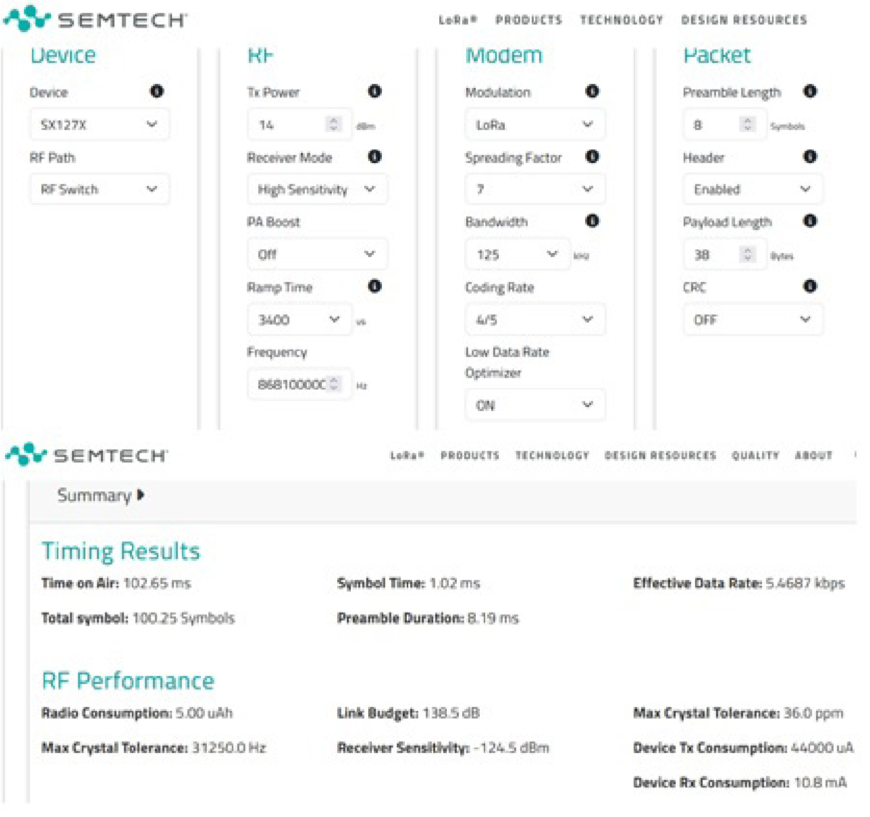


Figure 2. Data from the Semtech calculator

With the selected values of the individual parameters, the time for refreshing the data from the end devices is determined, subject to the restriction – the time for connecting to the network should be up to 30 seconds per day. The data from the end devices are published without receiving acknowledgement from the TTN server. To obtain data on the range of the gateway, in real operating conditions, the TTN Mapper platform was selected. It provides the ability to compile a map with the strength of the received signal from the end device at a specific point. In case the end device does not have a navigation receiver, the built-in receiver in mobile phones can be used. For this purpose, the TTN Mapper application for Android was used. Receiving the data in the TTN Mapper mobile application, their integration with the geographical coordinates of the relevant point and their transmission to the platform TTN Mapper is shown in Figure 3.

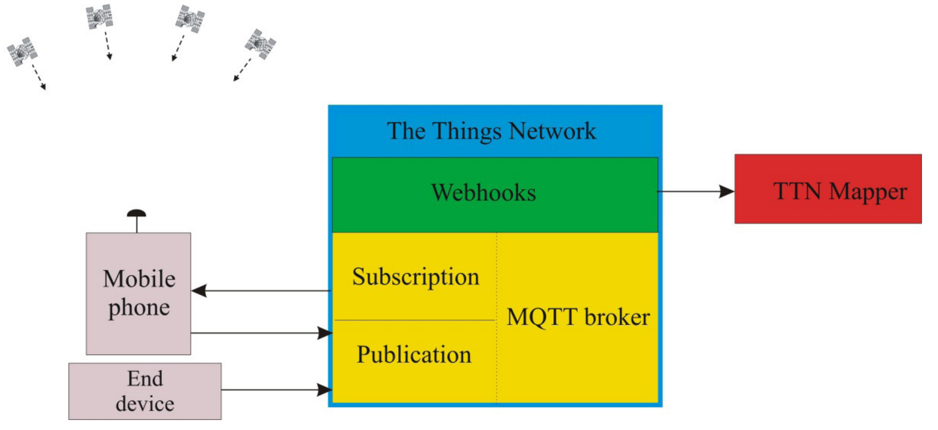


Figure 3. Block diagram of the connection between the individual platforms

To study the coverage, the end device with a temperature and humidity sensor was used. It was placed in close proximity to the phone with the TTN Mapper application installed. To receive the data from this device, the TTN Mapper application on the phone was programmed as a subscriber to the MQTT broker of the The Things Network platform. The metadata associated with the received messages contain the signal strength and connection quality received from the gateway. This metadata is integrated with the geographical coordinates of the corresponding point and sent to the network server. The TTN Mapper platform is connected to the TTN server using the Webhooks method. When data is received in the broker, it is immediately sent for visualization in TTN Mapper. The same data, if desired by the user, can be saved on the phone. It was chosen not to use the coverage visualization offered by TTN Mapper, but to use the data saved in the phone. They were further processed and stored as separate variables that were used to draw the figures. To study the influence of the antenna position, experiments were conducted in which the gateway was located outside a window, on the fourth floor, in a building of the VTU Todor Kableshkov. The gateway antenna is vertical to the ground surface, and the end device (ED) antenna is in a horizontal plane, in the first experiment. The direction of movement is counterclockwise from the location of the gateway. Its location and the locations where the terminal device was are shown in Fig. 4. The positions of the ED are recorded on the figure. The obtained data for the signal strength in the individual positions is shown in Fig. 5. The signal in the 30th position is the weakest, after it its value is less than the receiver sensitivity -124.5dBm (shown in Fig. 2). The distance between the gateway and position 30 is 297m in a straight line.

Figure 4. Position of the gateway and end device

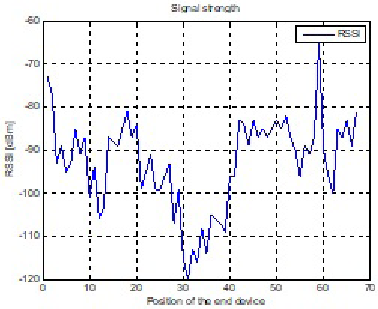


Figure 5. Signal strength at individual positions

A repeat attempt was made along the same route, with only the position of the ED antenna being changed, it was placed vertically relative to the earth's surface.

Figure 6. Position of the gateway and end device

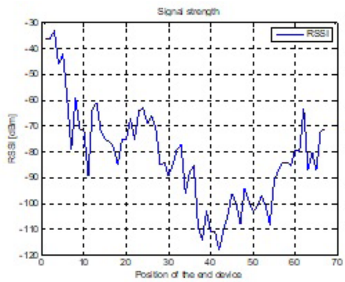


Figure 7. Signal strength at individual positions

The results are shown in Figure 6 and 7. It can be seen that in the area where there was no coverage before, the lowest RSSI value is now -118dBm. This provides a margin of 6.5dBm, with which a greater range of operation of the gateway can be achieved. The farthest distance in a straight line between the gateway and position 43 is 343m.

Conclusion

The projected LoRaWAN network makes it possible to implement a "smart irrigation system" with affordable materials, at a relatively low cost. Basic dependencies are presented for calculating network parameters necessary for determining the size of the payload and the time during which messages are sent to the network server. The range of the gateway has been tested only in an urban environment, with different polarizations of the antenna of the end device. The experiments were conducted using an antenna without directional action (with a circular diagram). It has been established that the range of action is greater when the polarization of the two antennas (of the ED and the gateway) is the same. In a specific application of the system (the location of the ED and the gateway is known), directional antennas with a higher gain can be used, thus a much greater range of action can be obtained. Taking into account the fact that the system is for use outside populated areas, the coverage will also be greater. The proposed system can be easily expanded by adding new ED, depending on specific needs.

REFERENCES

[1] ASHTON KEVIN, That 'Internet of Things' Thing, RFID Journal, 2009, ISSN 1936-7406, http://www.rfidjournal.com/article/view/4986.

[2] IEEE. IEEE Standard for Low-Rate Wireless Networks. IEEE Std 802.15.4-2024 (Revision of IEEE Std 802.15.4-2020), pp. 1 – 967. IEEE, 12 Dec. 2024. DOI: 10.1109/IEEESTD.2024.10794632.

[3] MOKO SMART, What is LPWAN (Low Power Wide Area Network). [online]. Available at: https://www.mokosmart.com/what-is-lpwan-low-power-wide-area-network [Accessed 9 May 2025].

[4] SILVA, M., MARQUES, D., VALADARES, D., FERREIRA, A., Assessment of LoRa Signal Strength in a Building Environment, Preprints 2024, 2024080663, https://doi.org/10.20944/preprints202408.0663.v1

[5] CALLEBAUT, G., LEENDERS G., BUYLE, C., CRUL, S., VAN DER PERRE, L., LoRa Physical Layer Evaluation for Point-to-Point Links and Coverage Measurements in Diverse Environments, in 2019 European Conference on Networks and Communications (EuCNC), Valencia, Spain, 2019, pp. 362 – 366, ISBN 978-1-7281-0306-8, doi: 10.48550/arXiv.1909.08300.

[6] SEYE, M., NGOM, B., GUEYE, B., DIALLO, M., A Study of LoRa Coverage: Range Evaluation and Channel Attenuation Model, 2018 1st International Conference on Smart Cities and Communities (SCCIC), Ouagadougou, Burkina Faso, 2018, pp. 1 – 4, ISBN 978-1-5386-5591-7, doi: 10.1109/SCCIC.2018.8584548.

[7] DE CAMARGO, E., SPANHOL, F., CASTRO E SOUZA, Á., Deployment of a LoRaWAN network and evaluation of tracking devices in the context of smart cities, Journal of Internet Services and Applications, 12(1), Article 8, pp. 1 – 17, ISSN 1869-0238, https://doi.org/10.1186/s13174-021-00138-7.

[8] BLENN, N., KUIPERS, F., LoRaWAN in the Wild: Measurements from The Things Network, arXiv preprint arXiv:1706.03086, pp. 1–10, ISSN 2331-8422, https://doi.org/10.48550/arXiv.1706.03086.

[9] SCHINSCHKE, J., SCHMIETENDORF, A., LoRaWAN Coverage Analysis in the Transportation Sector: A Real-World Approach, Engineering Proceedings,

2022, 27(1), 73, pp. 1 – 7. ISSN 2673-4591, https://doi.org/10.3390/ecsa-9-13321, https://doi.org/10.3390/ecsa-9-13321.

[10] ANGELOV K., N. MANCHEV, P., Kogias and S. Sadinov, Design and Development of a Platform for Test Applications in LoRa/LoRaWAN, Journal of Engineering Science and Technology Review (JESTR), Kavala Institute of Technology, 2019, pp. 17 – 21, ISSN 1791-9320, E-ISSN1791-2377.

[11] ANGELOV, K., MANCHEV, N., SADINOV, S., IVANOV, T., Planning and research of a radio coverage area in a LoRaWAN communication network, International Scientific Conference UNITECH 2020, November 20-21, 2020, Gabrovo, Volume I, pp. I-263-268, 2020, ISSN: 1313-230X.

[12] ALIPIO, M., BURES, M., Current testing and performance evaluation methodologies of LoRa and LoRaWAN in IoT applications: Classification, issues, and future directives, Internet of Things, Volume 25, April 2024, vol. 25, art. no. 101053, Apr. 2024, 35 pp., ISSN 2542-6605, doi: 10.1016/j.iot.2023.101053.

[13] MONTAGNY, S., LoRa – LoRaWAN and IoT, Savoie Mont Blanc University, 240 pp., ISBN 978-2-37741-084-2, https://www.univ-smb.fr/lorawan/en/free-book/.

[14] Semtech Corporation, "LoRa Calculator," Semtech Design Support, 2024. [Online]. Available: https://www.semtech.com/design-support/lora-calculator.

[15] Things Network, "The Things Network Documentation," 2024. [Online]. Available: thethingsnetwork.org. [Accessed: May-2024] https://www.thethingsnetwork.org/.

✉ **Prof. Dr. Emil Iontchev**
ORCID iD: 0009-0005-3585-7084
Department of Telecommunications and Safety Equipment and Systems
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: ionchev@vtu.bg

✉ **Rashko Vladimirov**
ORCID iD: 0009-0001-6600-4675
Department of Telecommunications and Safety Equipment and Systems
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: r.vladimirov90@gmail.com

✉ **Rosen Miletiev**
ORCID iD: 0000-0001-8050-1232
Technical University of Sofia
8, Kl. Ohridski Blvd., 1000 Sofia, Bulgaria
E-mail: miletiev@tu-sofia.bg

✉ **Tsvetan Petrov**
ORCID iD: 0009-0005-5637-3652
Department of Telecommunications and Safety Equipment and Systems
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: ceco_1@abv.bg

https://doi.org/10.53656/isct-2025.26

Electronics, Automation and Mechanization

A METHOD FOR BUILDING A TELECOMMUNICATION NETWORK TOPOLOGY BASED ON THE CLUSTER HIERARCHY

Illia Syvolovskyi
Ukrainian State University of Railway Transport, Kharkiv, Ukraine

Volodymyr Lysechko, Maksym Surhai
Olena Lozko, Serhiy Kaduk, Oleksii Mohylnyi
Ivan Kozhedub Kharkov National University of Air Forces, Kharkiv, Ukraine

Galina Cherneva
Todor Kableshkov University of Transport, Sofia, Bulgaria

Abstract. The growing complexity of distributed telecommunication environments, particularly in the context of the Internet of Things (IoT), edge computing, and real-time analytics, demands more adaptive and scalable network architectures. Traditional flat or static topologies are insufficient to support the high level of dynamism, heterogeneity, and fault tolerance required by modern systems. This article proposes a structured approach to the design of telecommunication network topologies based on hierarchical cluster organization, with the goal of optimizing data flow management and ensuring operational stability under dynamic conditions.

The suggested model incorporates a three-level architecture—Edge, Fog, and Cloud—each of which is responsible for distinct layers of data aggregation, processing, and transmission. Clusters within each layer are structured according to a star topology, enabling efficient coordination and minimizing the latency associated with inter-node communication. Central nodes within clusters manage subordinate nodes and serve as control points for load balancing and fault handling.

Special attention is given to the development of two key algorithms: one for the adaptive integration of new nodes based on metrics such as geographic proximity, task type, and cluster load; and another for automated topology reconfiguration in response to node failures or critical overloads. These algorithms include mechanisms for real-time monitoring, resource reservation, dynamic reassignment of tasks, and route optimization.

The proposed approach ensures enhanced fault tolerance, scalability, and adaptability of the network infrastructure. It is particularly suited for use in systems characterized by frequent structural changes and varying computational demands. The model has broad applicability in next-generation telecommunication systems, including distributed sensor networks, industrial IoT deployments, and intelligent infrastructure relying on Edge – Fog – Cloud paradigms.

Keywords: Telecommunication systems, cluster structure, load balancing traffic, data processing, optimization, distribution of data streams.

1. Introduction

The rapid expansion of smart devices, real-time services, and data-intensive applications has fundamentally transformed the landscape of modern telecommunication networks. These changes place increasing demands on network infrastructures, which must now provide not only high throughput but also flexibility, low latency, and fault tolerance in highly dynamic environments.

Conventional flat or centralized architectures often struggle to meet these challenges, particularly in systems where data sources are widely distributed, such as Internet of Things (IoT) ecosystems [5, 3, 7]. As a result, researchers and engineers have increasingly turned to hierarchical and decentralized models that can better align with the spatial and computational diversity of such environments.

Among the various approaches, multi-tier architectures that separate responsibilities between Edge [4], Fog [6], and Cloud layers have shown considerable promise. These models aim to balance local responsiveness with centralized efficiency. However, the design and management of such systems require robust structural frameworks capable of handling continuous node additions, workload fluctuations, and unexpected failures.

This paper addresses these challenges by introducing a method for constructing hierarchical cluster-based topologies. Rather than focusing solely on architectural design, the study emphasizes adaptive algorithms that enable real-time integration of new nodes and resilient reconfiguration in the face of faults. The goal is to provide a practical and scalable solution for distributed telecommunication infrastructures operating under unpredictable and dynamic conditions.

2. Hierarchical topology of clusters

In dynamic distributed telecommunication systems, a hierarchical network structure ensures stability, data transmission efficiency, and flexibility to changes. Localization of data transmission within clusters minimizes inter-cluster delays, and node redundancy guarantees system fault tolerance. The scalability of this structure makes it easy to add new nodes and clusters with minimal impact on the existing topology. At the same time, the hierarchical topology faces a number of challenges related to dynamic changes, load balancing, and fault tolerance. The main aspects of these challenges and possible solutions are summarized in Table 1.

Table 1. Challenges of hierarchical topology

Challenges	Solution
Dynamic changes in the network	Adaptation to changes in real time: adding nodes, rebuilding clusters.
Uneven load balancing	Centralized management of master nodes in clusters, automatic task reassignment.
Fault tolerance	Resource reservation, task redirection between nodes, traffic redistribution.

The proposed structure of the hierarchical cluster topology is based on the interaction of nodes within and between local clusters, which allows for effective traffic localization, increased fault tolerance, and flexibility to changes. The main levels of the hierarchy (Edge, Fog, Cloud) perform specific functions, each of which optimizes data processing and transmission within its area of responsibility. Fig. 1 shows a diagram of the hierarchical topology of clusters, which combines three main components: cluster structure, intercluster and interlevel interaction.

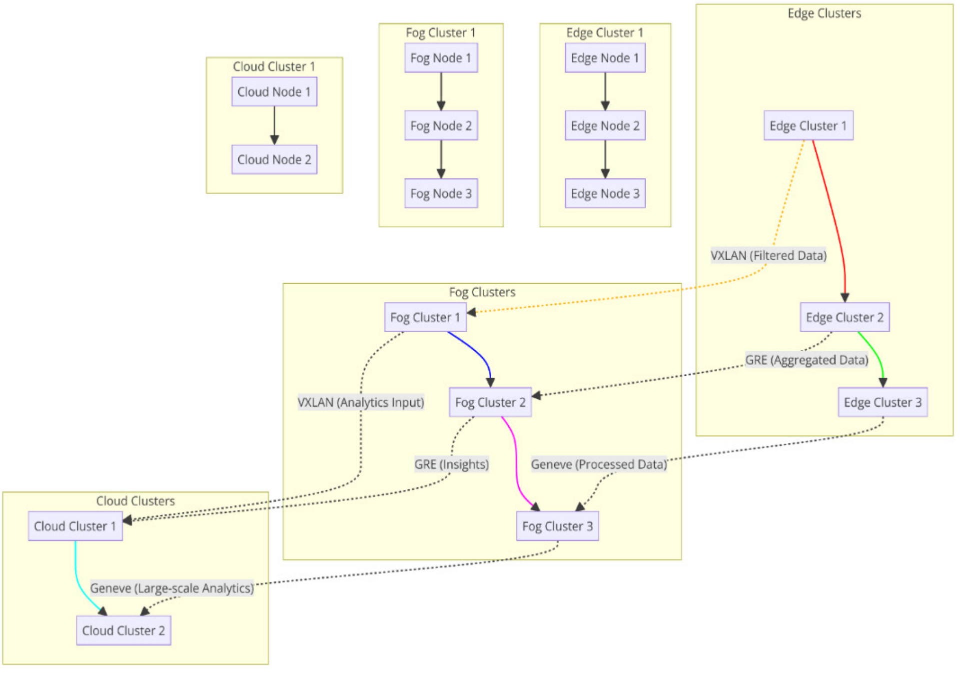


Figure 1. Diagram of the hierarchical topology of clusters

1. Structure of clusters. Each cluster is organized according to the "star" principle:

– The master node (centroid) is responsible for traffic management and coordination of the work of the slave nodes;

– The subordinate nodes perform specific computing tasks, transmitting the results to the master node.

2. Intercluster interaction:

– Clusters are connected through master nodes that form a star-to-star hierarchy.

– Traffic localization reduces the number of inter-cluster connections, reducing overhead.

3. Cross-tier interaction.

– Edge clusters process data directly from IoT devices, performing filtering and aggregation.

– Fog clusters perform pre-processing, providing balancing and optimization before transferring to the Cloud.

– Cloud clusters perform in-depth analysis and storage of large amounts of data.

3. Algorithm for adaptive addition of nodes to the system in the face of dynamic changes

To ensure the flexibility of the hierarchical topology, it is necessary to create a mechanism that allows you to effectively integrate new nodes into the existing system without affecting network performance [1, 2]. The algorithm for adaptive node addition includes stages.

Stage I. New node detection. When a new node is connected, the system registers its computing resources (CPU, RAM), power consumption, and communication channel bandwidth.

Stage II. Cluster compliance assessment. The most optimal cluster for the node is determined based on the following parameters: – geographical proximity; – cluster load; – type of tasks performed by the node.

Stage III. Checking the load balance. Before adding a node, the system evaluates the workload of the cluster head node. If the threshold value is exceeded, a decision is made to create a new cluster.

Stage IV. Topology update. After a node is added, the matrix of connections between nodes is updated, and data transmission routes are reconfigured.

V stage. System validation. The correctness of the node integration is checked, as well as the impact on the overall system performance. The block diagram of the algorithm is shown in Fig. 2.

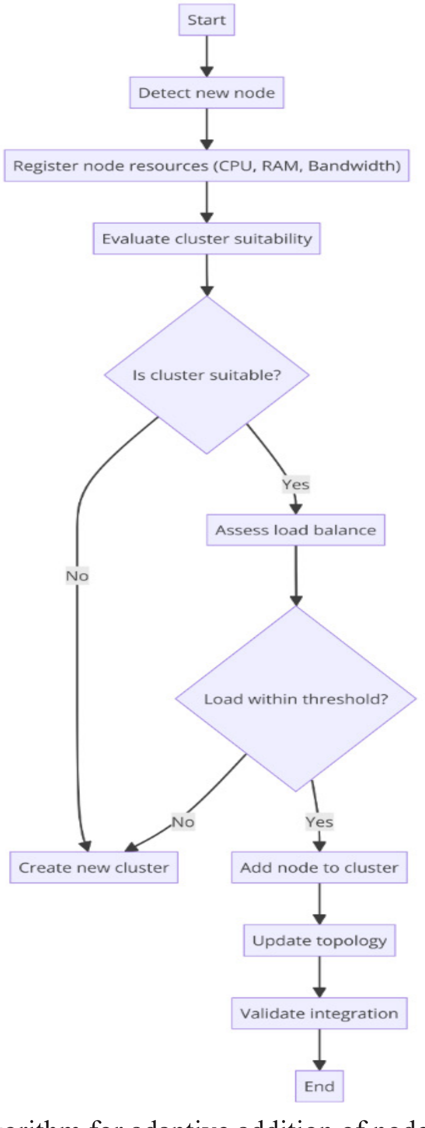


Figure 2. Algorithm for adaptive addition of nodes to the system

4. Algorithm for rebuilding the cluster structure during node failures or load surges

In case of node failure or a sharp increase in load, the system must adapt its topology to ensure stable operation [9]. The rebuilding algorithm includes (Fig. 3).

Stage I. Monitoring the state of nodes. The master nodes of clusters periodically check the status of the slave nodes (availability, resources).

Stage II. Identification of a problem situation. One of the critical situations is recorded: - failure of the master node; - failure of a slave node; - exceeding the permissible load.

Stage III. Reservation or reassignment of nodes. In case of failure of the master node, its functions are transferred to the backup node, which is automatically activated.

Stage IV. Redistribution of tasks. Tasks from the overloaded cluster are distributed among other clusters.

Stage V. Routing update. Routes between nodes are rebuilt to ensure efficient traffic.

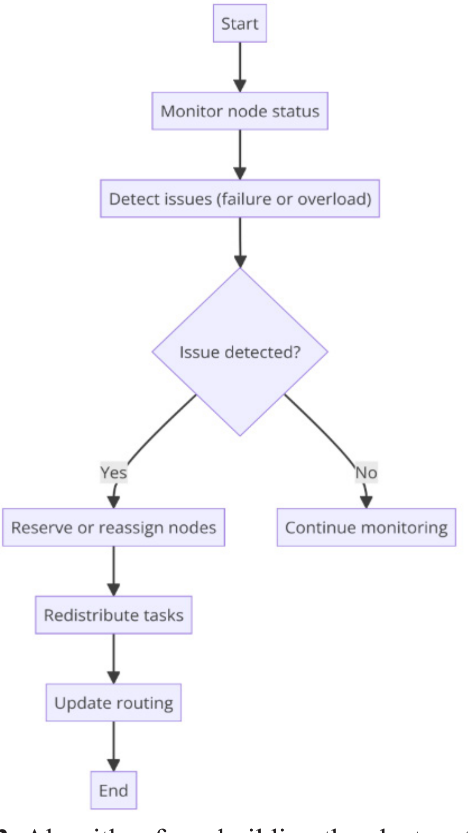


Figure 3. Algorithm for rebuilding the cluster structure

Conclusions

This paper presents a comprehensive approach to the design and management of hierarchical cluster topologies in distributed telecommunication networks, addressing the pressing need for scalable, adaptive, and fault-tolerant infrastructures in dynamic environments. The proposed model is based on a three-level architecture comprising Edge, Fog, and Cloud layers, where each level is responsible for specific tasks related to data processing, routing, and aggregation, depending on its proximity to data sources and computational capacity.

The core of the model is the organization of network nodes into clusters with star topology, which facilitates effective traffic coordination, localizes data transmission, and enhances fault isolation. The hierarchical structure simplifies system scalability, allowing seamless integration of new nodes and clusters with minimal impact on the existing configuration. Furthermore, the interaction between clusters and across tiers ensures balanced distribution of computational loads and optimized data flow throughout the network.

To support real-time adaptability, two key algorithms were developed: one for the adaptive addition of nodes, and another for restructuring the topology in the event of node failures or traffic surges. These algorithms incorporate mechanisms such as real-time monitoring, resource reservation, reassignment of roles, and route optimization, thereby ensuring stable operation under fluctuating workloads and failure conditions.

The proposed approach is particularly applicable to modern telecommunication systems, including IoT infrastructures, sensor networks, and intelligent control systems, where performance, reliability, and responsiveness are critical [1, 2, 8]. Future research may focus on the integration of machine learning techniques for predictive load balancing, energy-aware routing, and autonomous cluster formation to further enhance system intelligence and self-management capabilities.

REFERENCES

[1] ABDALLAH, W., VAL, T., Genetic-Voronoi algorithm for coverage of IoT data collection networks, 30th International Conference on Computer Theory and Applications (ICCTA), Alexandria, Egypt, 2020, pp. 16 – 22. doi: https://doi.org/10.48550/arXiv.2202.13735.

[2] GOUDARZI, M., ZAMANI, M., BUYYA, R., RANJAN, R., Distributed Decision Making for Efficient and Reliable Resource Management in Fog Computing Environments, IEEE Transactions on Computers, vol. 71, no. 1, pp. 141 – 155, 2022. https://doi.org/10.1109/TC.2021.3090918.

[3] DEB, K., ABOUHAWWASH, M., An optimality theory-based proximity measure for set-based multiobjective optimization, IEEE Transactions on Evolutionary Computation, 20(4), 2016, pp. 515 – 528. doi:10.1109/TEVC.2015.2483590.

[4] SATYANARAYANAN, M., The Emergence of Edge Computing, Computer, vol. 50, no. 1, pp. 30 – 39, Jan. 2017. https://doi.org/10.1109/MC.2017.9.
[5] GUERRERO, C., LERA, I.,·JUIZ, C., Genetic algorithm for multi-objective optimization of container allocation in cloud architectures, Journal of Grid Computing, 2018, vol. 16, no. 1, pp. 113 – 135, ISSN 1570-7873. doi: 10.1007/s10723-017-9419-x.
[6] YI, S., QIN, Z., LI, Q., Security and Privacy Issues of Fog Computing: A Survey, International Conference on Wireless Algorithms, Systems, and Applications (WASA), 2015, pp. 685 – 695. doi: 10.1007/978-3-319-21837-3_63.
[7] SKARLAT, O., NARDELLI, M., SCHULTE, S., BORKOWSKI, M., LEITNER, P., Optimized IoT service placement in the fog. Special Issue Paper. Open access. SOCA. Volume 11, 2017, pages 427 – 443. https://doi.org/10.1007/s11761-017-0219-8.
[8] ZHANG, Q., ZHU, Q., BASAR, T., Distributed Fog Computing in IoT for Sustainable Smart Cities: A Game-Theoretic Perspective, IEEE Internet of Things Journal, 2019, vol. 6, no. 5, pp. 7635 – 7647, ISSN 2327-4662. DOI: 10.1109/JIOT.2019.2913733.
[9] TRIVEDI, K., SAHNER, R., SHARPE at the Age of Twenty Two. CM SIGMETRICS Performance Evaluation Review, 2009, vol. 36, no. 4, pp. 52 – 57. ISSN 0163-5999. doi: 10.1145/1530873.1530882.

✉ **Illia Syvolovskyi**
ORCID iD: 0000-0002-4592-0965
Ukrainian State University of Railway Transport
Oboronny Val square, Ukraine
E-mail: ilyasvl95@gmail.com

✉ **Prof. Volodymyr Lysechko, DSc.**
ORCID iD: 0000-0002-1520-9515
Ivan Kozhedub Kharkov National University of Air Forces
77/79, Sumska St., Kharkov, Ukraine
E-mail: lysechkov@ukr.net

✉ **Dr. Maksym Surhai**
ORCID iD: 0000-0002-3979-005X
Ivan Kozhedub Kharkov National University of Air Forces
77/79, Sumska St., Kharkov, Ukraine
E-mail: maksimsurgaj@gmail.com

✉ **Olena Lozko**
ORCID iD: 0000-0002-6442-019X
Ivan Kozhedub Kharkov National University of Air Forces
77/79, Sumska St., Kharkov, Ukraine
E-mail: vladimirovnae952@gmail.com

✉ **Serhiy Kaduk**
ORCID iD: 0000-0002-2932-7416
Ivan Kozhedub Kharkov National University of Air Forces
77/79, Sumska St., Kharkov, Ukraine
E-mail: Kaskad434s@gmail.com

✉ **Oleksii Mohylnyi**
ORCID iD: 0009-0009-0238-9084
Ivan Kozhedub Kharkov National University of Air Forces
77/79, Sumska St., Kharkov, Ukraine
E-mail: alloexey@gmail.com

✉ **Galina Cherneva**
ORCID iD: 0000-0001-7441-0270
Department of Electrical Engineering and Physics
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: cherneva@vtu.bg

https://doi.org/10.53656/isct-2025.27

Materials Science and Technologies

EFFECT OF ZIRCONIUM SUBSTITUTION ON THE MICROSTRUCTURE AND FUNCTIONAL PROPERTIES OF INTERMETALLIC ALLOYS FOR HYDROGEN STORAGE

Nikolay Tonchev
Todor Kableshkov Transport University, Sofia, Bulgaria

Abstract. This study investigates hydrogen storage intermetallic alloys capable of reversibly absorbing and releasing hydrogen in the solid state. The focus is on optimizing the balance between hydrogenation/dehydrogenation kinetics and storage capacity. The effect of zirconium (Zr) substitution on the phase structure and functional properties of two base alloys – $Ti_{50}V_{25}Cr_{25}$ and Ti_2CrV – is analyzed. Initially exhibiting a single-phase BCC structure, these alloys develop a dual-phase structure (BCC + C15 Laves) upon Zr addition. Increasing Zr content enhances the fraction of the C15 phase, leading to improved kinetics and easier activation at room temperature, but also to a reduction in maximum hydrogen storage capacity. The study highlights the importance of phase evolution and structural modifications in achieving enhanced functional performance of hydrogen storage alloys.

Keywords: hydrogen storage alloys; zirconium; intermetallic compounds; hydrogenation kinetics; storage capacity, room-temperature activation

1. Introduction

Hydrogen storage is essential for achieving a sustainable energy transition and for integrating hydrogen as a clean energy carrier. Current research considers four main storage approaches, each of which has its own advantages and limitations, as well as its own potential for scaling up.

The most common method is gas compression [1, 2], whereby hydrogen is stored at high pressure (350 – 700 bar) in dedicated tanks. Despite its widespread use, particularly in transport, this method incurs high energy costs for compression and poses safety risks due to the high pressure. Additionally, it has limited volumetric efficiency.

Cryogenic storage involves liquefying hydrogen at extremely low temperatures (-253°C), which increases its energy density to ~70 g/L compared to its gaseous form. Significant energy losses from evaporation (up to 1% per day) and high capital costs for cryogenic tanks and insulation are the main challenges here [3].

Against this background, solid-state storage using metal hydrides and intermetallic alloys emerges as a promising alternative [4]. It offers high safety standards, good volumetric efficiency, and the ability to operate at moderate temperatures and pressures. Of the various approaches to hydrogen storage, solid hydrides are notable for their high volumetric density, safety, and potential for reversible operation under moderate conditions. Examples include metallic hydrides such as $LaNi_5$ and complex hydrides such as MgH_2 and $NaAlH_4$. However, these often suffer from slow kinetics and high dehydrogenation temperatures. Recent research on high-entropy hydrides (HEH) has revealed new ways to overcome these limitations, but challenges remain regarding the stability and cost-effectiveness of the materials.

In this context, the present study aims to analyse the role of zirconium substitution in the evolution of intermetallic hydrogen storage alloys, emphasising its effects on phase structure, hydrogenation kinetics, and storage capacity. It traces how the addition of zirconium contributes to the formation of new crystalline phases, such as C15 Laves, and how it facilitates activation at room temperature and improves the functional characteristics of the alloys. The study also presents a systematic approach to designing multi-element alloys through microstructural engineering, with the aim of developing robust, efficient, and scalable solutions for solid-state hydrogen storage.

2. Evolution and features of hydrogen storage alloys

The development of these alloys follows a consistent evolutionary logic, with each new generation of materials overcoming the limitations of its predecessors. Figure 1 illustrates the technological advances and milestones in this field.

Figure 1. Milestones in technological advances in hydrogen storage alloys

Traditional hydrogen-storing alloys of the AB_2, AB_3 and AB_5 types, which were developed between the 1960s and 1980s, formed the basis for modern designs of such materials. However, these alloys have limited hydrogen storage capacity and rapidly degrade in performance under cyclic operation (hydrogen loading and release). Research into AB_2 and AB_3 type alloys intensified in the late 1970s and 1980s [5, 6].

Ti-Cr-V series alloys represent a significant improvement on previous generations of hydrogen storage materials, demonstrating better performance at moderate temperatures and pressures. However, they still suffer from long incubation times before the first hydrogenation process and the presence of a low-pressure plateau during the initial absorption of hydrogen. These alloys emerged in the 1980s as an alternative to rare earth systems, undergoing continuous development and optimisation throughout the 1990s and 2000s [7].

Ti-Cr-V systems, along with alloys such as $TiMn_2$, remain the focus of intensive research aimed at enhancing kinetics and cyclic stability. Since 2020, there has been increased scientific interest and progress in this area [8].

Ti-Cr-V alloys innovate the microstructure by transitioning from a single-phase body-centred cubic (BCC) structure to a two-phase system including a BCC and a C15 Laves phase. At higher zirconium content (x = 2.0), an additional Zr-enriched phase is present. Zirconium substitution occurs in this generation of $Ti_{2-x}Zr_xCrV$ alloys, which is the subject of the present study [9].

The addition of zirconium resulted in a significant improvement in performance, expressed by faster initial hydrogenation and substantially reduced incubation time. However, this requires a trade-off - as the Zr content increases, there is a decrease in hydrogen storage capacity. The microstructure proves to be a crucial factor for the absorption efficiency: substitution with Zr leads to the formation of a C15 Laves phase alongside the BCC structure, which facilitates hydrogen diffusion and shortens the incubation time. For example, the time to reach 90% of the maximum capacity can be reduced from almost 3000 seconds to only 120 seconds at higher Zr content [10].

Additionally, nanocrystalline and fine-grained microstructures significantly accelerate the absorption process. The introduction of zirconium initiated the development of multi-element alloys (MEA), which are the focus of modern research. In the period 2024 – 2025, new alloy systems with improved microstructure and kinetics are reported [11, 12]. Although systematic studies on zirconium substitution and microstructural analysis started around 2010, it is the recent years that have marked significant progress in this direction.

The addition of zirconium resulted in a significant improvement in performance, expressed by faster initial hydrogenation and substantially reduced incubation time. However, this requires a trade-off – as the Zr content increases, there is a decrease in hydrogen storage capacity. The microstructure proves to be a crucial factor for the absorption efficiency: substitution with Zr leads to the formation of a C15 Laves phase alongside the BCC structure, which facilitates hydrogen diffusion and shortens the incubation time. For example, the time to reach 90% of the maximum capacity can be reduced from almost 3000 seconds to only 120 seconds at higher Zr content [10].

Additionally, nanocrystalline and fine-grained microstructures significantly accelerate the absorption process. The introduction of zirconium initiated the development of multi-element alloys (MEA), which are the focus of modern research. In the period 2024 – 2025, new alloy systems with improved microstructure and kinetics are reported [11, 12]. Although systematic studies on zirconium substitution and microstructural analysis started around 2010, it is the recent years that have marked significant progress in this direction.

Fourth-generation $TiMn_2$-based alloys have been known since the 1970s, but their active development through the incorporation of zirconium (Zr) to create multicomponent systems only began after the year 2000 and continues to this day [13]. These alloys are characterised by a single-phase C14 Laves structure. The addition of Zr increases the lattice parameters and unit cell volume. This structural expansion improves the kinetics of hydrogen absorption and desorption, leading to better practical storage efficiency. The main challenges with these alloys relate to synthesis complexity, phase composition control, and potential capacity reduction at high zirconium content.

The latest generation of high-performance alloys for hydrogen storage includes nanostructured multi-element alloys (MEAs), such as $Ti_{21}Zr_{21}Fe_{41}Ni_{17}$. These alloys belong to the intermediate-entropy and multi-element class of systems, which are dominated by the C14 Laves phase and contain a small amount of the BCC phase. These alloys are characterised by rapid hydrogenation kinetics and the ability to reversibly store hydrogen at room temperature. They also have coherent nanograin boundaries that facilitate efficient hydrogen diffusion.

Despite their promising properties, these alloys have two main drawbacks: the complexity and high cost of synthesis, and the fact that their industrial scalability has not yet been fully demonstrated.

High-entropy alloys (HEAs) with a nanocrystalline structure exhibit extremely fast hydrogenation kinetics, reaching 90% saturation in around 100 seconds at room temperature – approximately five times faster than conventional body-centred cubic (BCC) alloys [14]. This is due to the high density of grain boundaries and defects in the nanostructure, which provide multiple diffusion pathways and active sites for hydrogen nucleation. Further improvements can be achieved by introducing defects or refining the microstructure using methods such as ball milling, or by adding catalysts. For instance, TiFe-based alloys treated with transition metals (V, Ni and Pd) demonstrate a notable decrease in particle size, an increased surface area, and the formation of new catalytically active centres. This results in accelerated hydrogen absorption. Some of these composites reached over 96% saturation in only 100 seconds [15].

In conclusion, a review of the development of hydrogen storage alloys reveals a clear evolution from classical AB_5 and Ti-Cr-V systems to modern, multi-element, high-entropy alloys with an advanced microstructure. Modern $TiMn_2$ alloys containing Zr, as well as MEA/HEA systems, demonstrate high energy efficiency during absorption and desorption cycles thanks to their fast kinetics, low activation energy and excellent cyclic stability.

Optimisation of water-scavenging alloys and the effect of zirconium addition.

The main objective in developing new hydrogen-conserving alloys is to increase their energy efficiency during absorption and desorption cycles. This includes accelerating kinetics, reducing activation energy, and ensuring stability during repeated use.

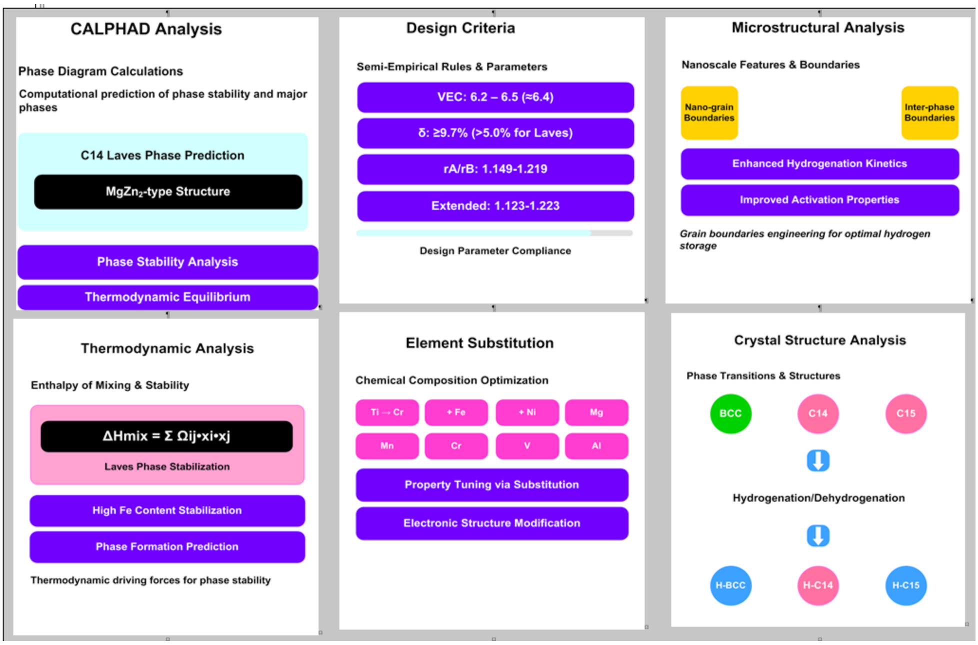


Figure 2. Process flow diagram for designing a hydrogen-containing alloy

Fig. 2 presents a workflow diagram of the design of such alloys, where each step is directed towards a specific goal.

The optimisation process begins with the application of semi-empirical criteria, such as valence electron concentration (VEC), atomic size mismatch (δ), and radius ratio. The phase stability of the alloys is then predicted using thermodynamic calculations and the CALPHAD method. Particular attention is paid to the microstructure, since nanograin structures and interfacial boundaries can greatly enhance hydrogenation kinetics. The composition of the alloys is tailored to the specific application by monitoring the effect of replacing titanium (Ti) with zirconium (Zr) or adding iron (Fe) or nickel (Ni). The enthalpy of mixing (ΔH_(mix)) is also employed to evaluate thermodynamic stability, particularly for C15 Laves phases that are rich in iron (Fe).

One of the most striking examples of effective optimisation is the addition of zirconium (Zr) to $Ti_{50}V_{25}Cr_{25}$ and Ti_2CrV alloys. Although these alloys have good potential, they suffer from slow activation and low initial kinetics, which limits their application, particularly in transport. Adding Zr creates a second crystalline C15 Laves phase, improving kinetics and facilitating activation at room temperature [10, 11].

For $Ti_{20-x}Zr_xCrV$ alloys with x = 0.5 – 2.0, significant improvements in activatability are observed: absorption begins immediately at room temperature and pressure of around 20 atmospheres. With the lowest Zr content (x = 0.5), a short incubation period is evident; however, at higher values (x ≥ 1.0), this disappears completely. This is a significant advantage over the pure Ti_2CrV alloy, which requires preheating to 400 °C for activation [10].

Conversely, a decrease in hydrogen storage capacity is observed as Zr content increases, from 4.4 wt.% at x = 0 to 2.3 wt.% at x = 2.0. This is due to the larger atomic mass of zirconium and the lower hydrogen-binding capacity of the C15 phase. A similar trend is observed for $Ti_{50-x}V_{25}Cr_{25}Zr_x$ alloys, where the capacity initially increases to 2.7 wt% at x = 7, but then decreases to 2.27 wt% at x = 9 due to the dominance of the C15 phase [10].

Furthermore, adding Zr improves desorption kinetics, shortening the time taken to reach equilibrium upon hydrogen release and increasing the desorption ratio (e.g. up to 47% for $Ti_{43}V_{25}Cr_{25}Zr_7$ versus 43% for an alloy without Zr) [10].

In summary, zirconium substitution has been shown to be an effective method of optimising hydrogen-conserving alloys. While it improves kinetics and facilitates activation, careful balancing with respect to storage capacity is required. This highlights the need for a systematic approach to designing new alloys that combines microstructural control, thermodynamic stability, and functional efficiency.

Conclusion

Zirconium substitution and nanostructuring are emerging as key strategies to increase efficiency, but they require careful balancing to avoid a significant reduction in storage capacity. Optimising these parameters in the latest generation of alloys lays the foundation for developing reliable, safe, and scalable solid-state hydrogen storage solutions that can meet the demands of energy applications. From here, more important conclusions can be drawn Zirconium (Zr) significantly improves hydrogen storage efficiency by tailoring phase composition, lattice parameters, and binding energies.

1. Moderate additions of zirconium (Zr) (5 – 15 at%) optimise the kinetics, thermodynamics, and cycle stability of Zr-modified alloys, making them superior to conventional systems such as $LaNi_5$ and TiFe.

2. However, excessive Zr reduces storage capacity, although it remains useful for applications that prioritise rapid hydrogen cycling over maximum storage.

3. Zirconium alloying influences phase transformations during hydrogen uptake by stabilising secondary phases (Laves phases/hydrides), accelerating kinetics through refined microstructures and modulating thermodynamic stability.

4. The $Ti_{43}V_{25}Cr_{25}Zr_7$ alloy is an example of optimum zirconia content, demonstrating rapid hydrogen uptake (230 s) while maintaining cyclic stability.

5. In nuclear alloys, zirconium (Zr) plays a dual role in hydride formation and α/β phase shifts, highlighting its impact on performance and degradation.

REFERENCES

[1] RIVARD, E., TRUDEAU, M., ZAGHIB, K., Hydrogen Storage for Mobility: A Review. Materials (Basel). 2019 Jun 19;12(12):1973. doi: 10.3390/ma12121973. PMID: 31248099; PMCID: PMC6630991.

[2] ABDIN, Z., TANG, C., LIU, Y., CATCHPOLE, K., Large-Scale Stationary Hydrogen Storage Via Liquid Organic Hydrogen Carriers. Iscience. 2021 AUG 9;24(9):102966. doi: 10.1016/J.ISCI.2021.102966. PMID: 34466789; PMCID: PMC8382998.

[3] RAMPAI, M., MTSHALI, C., SEROKA, N., KHOTSENG, L. Hydrogen Production, Storage, and Transportation: Recent Advances. RSC ADV. 2024 FEB 23;14(10):6699-6718. doi: 10.1039/D3RA08305E. PMID: 38405074; PMCID: PMC10884891.

[4] SU, J., LIU. J.. Catalysis Process Development for Hydrogen Storage/Generation With LOHCS. Top Catal. 2021;64(7-8):469. doi: 10.1007/S11244-021-01466-5. EPUB 2021 JUL 2. PMID: 34230789; PMCID: PMC8250554.

[5] VOLODIN, A., DENYS, R., WAN, C., WIJAYANTI, I., SUWARNO, TARASOV, B., ANTONOV, V., YARTYS, V., Study of Hydrogen Storage and Electrochemical Properties of AB2-Type TI0.15ZR0.85LA0.03NI1.2MN0.7V0.12FE0.12 Alloy, Journal of Alloys and Compounds, Volume 793, 2019, pages 564 – 575, ISSN 0925-8388, https://doi.org/10.1016/j.jallcom.2019.03.134.

[6] KOHNO, T., YOSHIDA, H., KANDA, M., Hydrogen Storage Properties of LA(NI0.9M0.1)3 Alloys, Journal of Alloys and Compounds, Volume 363, ISSUES 1 – 2, 2004, pages 254 – 257, ISSN 0925-8388, https://doi.org/10.1016/s0925-8388(03)00429-8.

[7] KAZAKOV, A., BLINOV, D., BODIKOV, V., MITROKHIN, S., VOLODIN, A., Hydrogen Storage and Electrochemical Properties of Annealed Low-Co AB5

Type Intermetallic Compounds, International Journal of Hydrogen Energy, Volume 46, ISSUE 25, 2021, pages 13622 – 13631, ISSN 0360-3199, https://doi.org/10.1016/j.ijhydene.2020.12.071.

[8] LIU, J., SUN, L., YANG, J., GUO, D., CHEN, D., YANG, L., XIAO, P., TI-MN Hydrogen Storage Alloys: From Properties to Applications. RSC ADV. 2022 Dec 14;12(55):35744-35755. doi: 10.1039/D2RA07301C. PMID: 36545097; PMCID: PMC9748650.

[9] BANERJEE, S., RUZ, P., KUMAR, A., Evaluation of Ti_2CrV Alloy for Solid State Hydrogen Storage Applications. BARC Newsletter, Jan.–Feb. 2020, p. 31. ISSN 0971-3559.

[10] ZENG, Q., WANG, F., LI, Z., RONG, M., WANG, J., WANG, Z.,
Influence of ZR Addition on the Microstructure and Hydrogenation Kinetics of TI50-XV25CR25ZRX (X = 0, 5, 7, AND 9) Alloys. Materials (BASEL). 2024 Mar 16;17(6):1366. doi: 10.3390/MA17061366. PMID: 38541521; PMCID: PMC10972261.

[11] MONSALVE, D., ULATE-KOLITSKY, E, MARTÍNEZ-AMARIZ, A.-D,, HUOT, J., Effects of TI Substitution By ZR, on the Microstructure and Hydrogen Storage Properties of TI2-XZRXCRV (X = 0.5, 1, 1.5, 2) Alloys. Heliyon. 2024 Aug 3;10(15):E35739. doi: 10.1016/J.HELIYON.2024.E35739. PMID: 39170160; PMCID: PMC11336814.

[12] ANDRADE, G., EDALATI, P., DANGWAL, S., EDALATI, K., FLORIANO, R., Microstructural Characterization and Hydrogen Storage Properties at Room Temperature of TI21ZR21FE41NI17 Medium Entropy Alloy – appl. Energy mater.,2025, vol. 8, 4, doi: 10.1021/acsaem.4c02468.

[13] RUI, X., CHENG, T., LI, C., YANG, X., RONG, J., Properties of Ti-Based Hydrogen Storage Alloy, Journal of Power And Energy Engineering, vol.12 No.3, 2024.

[14] LUO, L., HAN, H., FENG, D,, LV, W., CHEN, L., LI, L., ZHAI, T., LIU, S., SUN, S., LI, Y., PEI, W., CUI, J., LI, Y., Nanocrystalline High Entropy Alloys With Ultrafast Kinetics and High Storage Capacity for Large-Scale Room-Temperature-Applicable Hydrogen Storage, Renewables, 2(1), 2024, pp. 24 – 41, ISSN 2751-2401 (Online). doi 10.1002/rew2.15.

[15] HAN, Z., ZHAI, T., YUAN, Z., SUN, H., LI, T., ZHANG, L., ZHENG, C., ZHANG, Y., Influence on Microstructure and Hydrogen Storage Properties of Tife-based composites with transition metals v, ni, and pd by mechanical milling, SSRN Electronic Journal, 2024, ISSN 1556-5068, doi: 10.2139/ssrn.4803425.

✉ **Prof. Nikolay Tontchev, DSc.**
ORCID iD: (0000-0002-9110-2488)
Department of Machine Elements and Chemistry
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: tontchev@vtu.bg

https://doi.org/10.53656/isct-2025.28

Materials Science and Technologies

MULTIPARAMETRIC STUDY OF CUTTING ACCURACY IN LASER CUTTING BY CONTROLLING TECHNOLOGICAL PARAMETERS

Nikolay Hristov, Victor Georgiev
Todor Kableshkov University of Transport, Sofia, Bulgaria

Abstract. The article presents a comprehensive study on optimizing laser cutting processes by analyzing the influence of key technological parameters - cutting speed, laser power, and auxiliary gas pressure -on the accuracy and quality of cuts in C235 steel. The research employs a Design of Experiments (DOE) methodology, specifically the Extended Central Composition design, to systematically investigate how these parameters affect geometric precision and the heat-affected zone (HAZ) on both sides of the laser cut.

The study begins by outlining the fundamentals of laser cutting, emphasizing the importance of controlling process parameters to achieve high-quality results. It highlights that while previous research has focused on cut quality, few have addressed the accuracy of specific geometric shapes. The authors aim to fill this gap by developing a methodology that minimizes experimental trials while maximizing insight into parameter effects.

Using a Durma HD-F 4020/4KW fiber laser, the researchers conducted experiments on 3mm thick C235 steel samples shaped into 10mm squares. Measurements of diagonal deviations and HAZ widths were taken from both the upper and lower sides of the cuts. The DEFMOT method, developed by the authors, was used to analyze the differences in shape accuracy and thermal effects.

Regression models were derived to quantify the relationship between input parameters and output quality indicators. The results showed that increasing gas pressure and cutting speed generally led to greater differences in accuracy between the two sides of the cut. However, the influence of laser power was less consistent. The models demonstrated high correlation coefficients ($R > 0.95$), indicating strong predictive capabilities.

The study concludes that a multifactorial approach enables precise control over laser cutting outcomes. The developed methodology allows for the prediction and optimization of dimensional accuracy and thermal effects, contributing to more efficient and reliable laser cutting processes in industrial applications.

Keywords: Technological parameters of laser cutting; DOE; Laser quality

1. Introduction

In typical melting laser cutting, the laser energy melts the material and the assist gas jet removes the molten material [1, 2]. Therefore, less laser energy is required compared to cutting by evaporation [3]. The laser heat is combined with an exothermic chemical reaction with the auxiliary gas and provides the heat required to melt and evaporate the cut material. A number of fundamental studies have been carried out in the field of laser cutting [4], in which the quality has been investigated, without however affecting the accuracy of cutting a particular geometric shape. Nevertheless, the literature cited defines multi-criteria process parameters related to cut quality. Higher laser power generally increases the material removal rate, but can also increase the surface roughness if not properly controlled. It has been studied in the literature [5] that optimisation of parameters such as laser power, pulse frequency and cutting speed can minimise surface roughness and improve cut quality. In order to provide a solution for precision laser cutting of materials, it is necessary to control the accuracy and quality depending on the technological parameters. Such an approach provides solutions to achieve complex geometries with increased requirements. Although individual parameter optimisation would be required for each new target, geometry or material class, it is desirable to base the optimisation process on as few experiments as possible to determine the direction of the search. To search for such a methodology, a central-optimal compositional design was selected from the DOE with the response surface.

In fact, different types of laser systems already exist, including gas-assisted systems. Such systems offer many advantages, including high quality output. The selection of the most suitable process parameters becomes an important practical task to improve both the quality of the product and the productivity of laser cutting. Identification of individual control factors on geometrical and other parameters of the cut material is precisely the solution of the productivity problem as a compromise of a modern technological solution [6].

Design of Experiments (DOE) is a systematic method that has been applied to many optimisation problems. The Response Surface Method (RSM) [7].

The objective of the study is to compose a knowledge-based approach using specified parameters from the DOE of the laser cutting process to achieve the desired results in terms of a set of quality parameters. The approach demonstrates how controlling parametric combinations of cutting speed, auxiliary gas pressure, power and frequency leads to better surface quality of laser cut samples. The quality of the resulting cut is determined by pre-selected quality indicators such as geometric accuracy, cutting slope and temperature shock width. To achieve the objective, an experimental parametric analysis plan, Extended Central Composition (by Roland A. Fisher), is applied to laser-cut billets of elementary geometric shape of C235 steel. On the basis of the conducted experiment and the results of the study, a generalised methodology is created using the author's DEFMOT approach [8, 9].

With the help of the methodology it is necessary to analyse the deviation of quality parameters depending on the design of technological modes.

2. Problem statement and research and measurement methods.

In this section, first of all, the problem of the objective of the study - optimisation of laser cutting parameters in order to minimise the error in the slope accuracy in the dimensions and the width of the temperature zone of the cut - is defined, and the factors and levels of the input control parameters are identified.

The main factors influencing the quality of the cut in laser cutting are

– Cutting speed (S).

– Laser power (P).

– Auxiliary gas pressure (G).

The problem statement involves creating a 3D model of a sample, a square with a side size of 10 mm, in CAD software. This CAD model was then converted into a DXF file and fed into the laser cutting machine, Fig. 1.

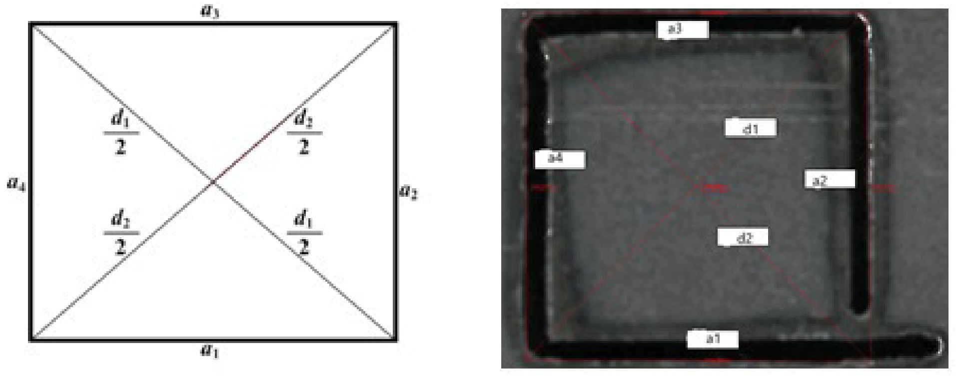


Figure 1. Significance of the modified ideal quadrant for different combinations of process factors compared to the CAD model

The equipment used is a Durma HD-F 4020/4KW fibre laser with a 3mm thick cut sheet. The object of study is the most widely used rolled steel C235. Its chemical composition is given in Table 1.

Table 1. Chemical composition of C235 steel

C	Si	Mn	Ni	S	P	Cr	N	Cu
≤0,22	≤0,05	≤0,60	≤0,30	≤0,040	≤0,040	≤0,30	≤0,012	≤0,30

The laser head has an integrated sensor system that monitors the cutting process and provides relevant information to the operator. In this way, it is possible to carry out the technological modes required to achieve the objectives of the study. The control modes of the experiments are changed according to the combinations indicated in Table 2, carefully for each different line of the plan. After the cutting process, the square specimens are removed from the sheet frame and the results of the diagonals on both sides described in Table 2 are obtained on the basis of the measurements for each combination on both sides of the contact surface. All values given in Table 2 are quantitative numerical measurements calculated on a pixel basis.

Table 2. Influence of different combinations of input process parameters on the investigated output parameters measured on both sides of the cut

№ upper side I	V, m/min	P, W	p, bar	d I_1	d I_2	HAZ I sr
1	2400	1700	11,130	15,741	15,680	0,370
2	3000	1700	11,172	15,741	15,860	0,365
3	2400	2500	11,215	15,920	15,890	0,391
4	3000	2500	11,172	15,860	15,890	0,391
5	2400	1700	11,172	15,861	15,771	0,391
6	3000	1700	11,130	15,861	15,771	0,339
7	2400	2500	11,172	15,861	15,800	0,391
8	3000	2500	11,257	15,891	15,981	0,360
9	3000	2100	11,257	15,860	15,861	0,344
10	2400	2100	11,215	15,890	15,890	0,397
11	2700	2500	11,215	15,950	15,890	0,391
12	2700	1700	11,172	15,920	15,771	0,391
13	2700	2100	11,299	15,980	15,950	0,365
14	2700	2100	11,299	15,980	15,920	0,365
№ down side II	**V mm/ min**	**P, W**	**p, bar**	**d II_1**	**d II_2**	**HAZ II sr**
1	2400	1700	11,046	15,741	15,892	0,571
2	3000	1700	11,088	15,651	15,920	0,608
3	2400	2500	11,215	15,890	16,011	0,656
4	3000	2500	11,172	15,771	15,981	0,693
5	2400	1700	11,088	15,651	15,801	0,603
6	3000	1700	11,046	15,621	15,801	0,651

7	2400	2500	11,130	15,801	15,861	0,656
8	3000	2500	11,215	15,950	15,863	0,682
9	3000	2100	11,172	15,771	15,981	0,667
10	2400	2100	11,215	15,831	15,980	0,608
11	2700	2500	11,299	15,950	16,040	0,672
12	2700	1700	11,130	15,741	15,920	0,640
13	2700	2100	11,130	15,741	16,010	0,698
14	2700	2100	11,299	15,860	16,100	0,677

From the measured diagonals, the difference in shape accuracy on both sides of the cut is determined using a DEFMOT method developed by the author. The details of determining the accuracy difference itself are defined by the algorithm shown in Figure 2. This algorithm determines the influence of individual process input factors on the accuracy difference.

3. Analysis of Results and Conclusions

Reliable design is a practical and rational method for optimising product qualities that are minimally sensitive to various causes. For this reason, an experiment was set up to study the combinations for the difference in accuracy on both sides of the cut that determine the best quality. The highly parametric influence of process parameters on the HAZ was evaluated. Since the laser beam speed and power and the cutting gas pressure were assumed to be the key parameters, regression models were derived for Δ - the difference in accuracy based on the algorithm shown in Fig. 2 and ΔHAZ – the difference in temperature influence zones on both sides of the cut with respect to these parameters.

To derive the regression models, the cutting process was repeated continuously for 9 modes with different control setting factors (those in Table 2). The difference between the upper and lower gauge dimensions and the dimensional error (Δ) are determined in the algorithm (Fig. 2). The derived polynomial models are of second order based on standard methodology [10]. All regression models have the form (1) and are derived for coded units:

(1) $y(h) = b_0 + \sum_{i=1}^{l} b_i x_i + \sum_{i=1}^{l-1}\sum_{j=i+1}^{l} b_{ij} x_i x_j + \sum_{i=1}^{l} b_{ii} x_i^2 + \varepsilon$

Setting the ideal CAD model – a cut square with side a = 10 mm and diagonals $d_1 = d_2$ = 14,142 mm

↓

Conducting experiments varying the technological parameters to establish the relative accuracy in the difference in dimensions on both sides of the cut made

↓

Measurement of the two real diagonals d_1 and d_2 to determine the two deviations from the accuracy of each of the diagonals for each of the sides marked with *I* and *II*

$$RI_{1\ 14,142} = \frac{dI_1 - 14{,}142}{14{,}142},\quad RI_{2\ 14,142} = \frac{dI_2 - 14{,}142}{14{,}142};$$

$$RII_{1\ 14,142} = \frac{dII_1 - 14{,}142}{14{,}142},\quad RII_{2\ 14,142} = \frac{dII_2 - 14{,}142}{14{,}142}$$

↓

Determine the relative deviation from shape accuracy for each side

$$\Delta RI = \frac{RI_1 + RI_2}{2} \cdot 100\%,\quad \Delta RII = \frac{RII_1 + RII_2}{2} \cdot 100\%$$

↓

Determine the difference in accuracy and derive a regression model for it

↓

Analyze the influence of the studied parameters on the difference in accuracy from the shape of the CAD model

Figure 2

where b are the coefficients of the regression model, xi, xj are the different control factors, l is the total number of control factors, y is the response and ε is the error.

When the data in Table 2 were processed, the following coefficients of model (1) were obtained and it took the following form for the two criteria under study.

Δ – the difference in accuracy on both sides of the cut

(2)

$$\Delta = -0.105769 + 0.159 X_1 - 0.141 X_2 + 0.352 X_3 - 0.04375 X_1 X_2 - 0.09375 X_1 X_3 - 0.108462 X_2^2 - 0.173750 X_2 X_3 + 0.306538 X_3^2$$

The multiple correlation coefficient was R =0.9533, and the calculated Fisher's F criterion of 6.2289 was greater than the tabulated one of 4.8183, (0.05, 8, 5).

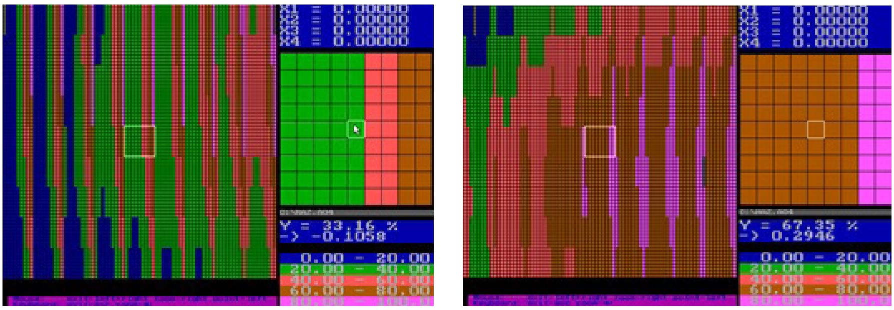
Δ HAZ – the difference between the temperature influence areas on both sides of the section

(3)

$$\Delta HAZ = 0.294625 + 0.0348 X_1 + 0.0218 X_2 + 0.0121 X_3 - 0.0276251 X_1^2 - 0.006 X_1 X_2 + 0.00975 X_1 X_3 - 0.029625 X_2^2 - 0.0075 X_2 X_3 + 0.027875 X_3^2$$

The coefficient of multiple correlation was R=0.9736 and the calculated Fisher's F criterion of 8.0949 was greater than the tabulated one of 5.9988, (0.05, 9, 4).

Immediately below the models are the results of their tests, at a selected significance level of 0.05, indicating that the statistical error is 5%.



Model (2) Model (3)

Figure 3. Graphical interpretation of the criteria studied using the illustration given in [9]

The images shown in Fig. 3 were generated using the author's DEFMOT analysis and optimisation system. This system allows automatic recalculation of the corresponding criterion depending on the control parameters. These calculations are shown graphically in Fig. 4 and Fig. 5.

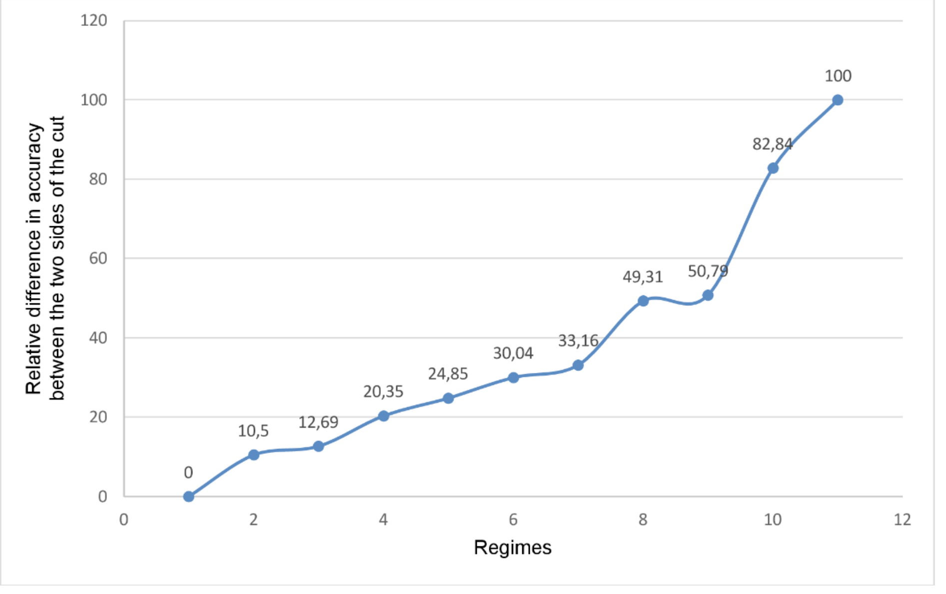


Figure 4. Variation, calculated as a percentage, of Δ, the difference in accuracy between the two sides of the cut, obtained for the regimes listed in Tab. 3.

Table 3. Regimes determining the parametric influence on the accuracy difference

	Regimes										
	1	2	3	4	5	6	7	8	9	10	11
%	0	10,5	12,69	20,35	24,85	30,4	31,16	49,31	50,79	82,84	100
V, mm/min	2400	2400	2400	2700	2700	2700	2700	2700	3000	3000	3000
P, W	1700	2500	2100	1700	2500	2100	2100	1700	2500	2100	1700
p, bar	0.4	0.4	0.4	0.4	0.4	0.4	0.5	0.6	0.6	0.6	0.6

From Fig. 4 and Table 3, it can be seen that the accuracy difference increases with increasing gas velocity and pressure, and in terms of power, the nature of its variation is unidentified and inconsistent.

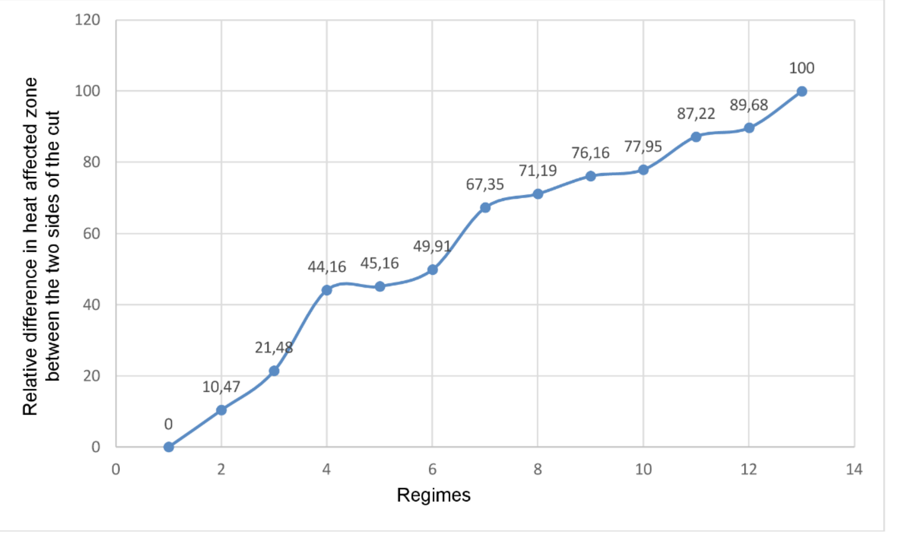


Figure 5. Variation, calculated as a percentage, of Δ HAZ, the difference of the heat affected zones on the two sides of the cut, obtained for the regimes listed in Table 4

Table 4. Regimes determining the parametric influence on the difference of the heat affected zones on both sides of the cut

	Regimes												
	1	2	3	4	5	6	7	8	9	10	11	12	13
%	0	10,47	21,45	44,16	45,16	49,91	67,35	71,19	76,16	77,95	87,22	89,68	100
V, mm/min	2400	2400	2400	2400	3000	2400	2700	3000	2700	3000	3000	2700	3000
P, W	1700	1700	1700	2500	1700	2500	2100	2500	2100	1700	2500	2100	1700
p, bar	0.5	0.4	0.6	0.6	0.4	0.4	0.5	0.4	0.4	0.6	0.6	0.6	0.6

From Table 4 it can be seen that no direct correlation can be made between the parameters studied. The model (3) can be used for a specific shape for the steel and thickness studied.

Conclusions

In a planned multifactorial and two-criteria experiment, the influence of the three main technological parameters on the accuracy of the cut contour, measured simultaneously on both sides of the cut, was determined. The steel studied was C235 with a thickness of 3 mm. The study clarified how, by controlling the parametric combinations (such as cutting speed and power, and auxiliary gas pressure), it is possible to predict a better quality in terms of accuracy of laser cut specimens, when both sides of the laser cut are controlled simultaneously. The effects of selected process parameters of the considered process on pre-selected quality indicators were evaluated, and a methodology to optimise the dimensional accuracy on two targets by minimising a function, i.e. the dimensional error, was proposed.

It was found that increasing the gas velocity and pressure increases the relative accuracy difference on both sides of the cut. It was found that there was not a complete correlation between the selected process factors and Δ HAZ, the difference in temperature influence zones on both sides of the cut.

REFERENCES

[1]UHLMANN E., MULLANY, B., BIERMANN, D., RAJURKAR, K., HAUSOTTE, T., BRINKSMEIER, E., Process chains for high-precision components with micro-scale features, CIRP Annals, Volume 65, Issue 2, 2016, pp. 549 – 572, https://doi.org/10.1016/j.cirp.2016.05.001.

[2] FEISTLE M., GOLLE, R., VOLK, W., Determining the Influence of Shear Cutting Parameters on the Edge Cracking Susceptibility of High-strength-steels Using the Edge-fracture-tensile-test, Procedia CIRP, Volume 41, 2016, pp. 1078 – 1083, https://doi.org/10.1016/j.procir.2016.01.007.

[3] SHARMA A., YADAVA, V., Experimental analysis of Nd-YAG laser cutting of sheet materials – A review, Optics & Laser Technology, Vol. 98, 2018, pp. 264 – 280, https://doi.org/10.1016/j.optlastec.2017.08.002.

[4] YI, P.; LIU, Y. C., SHI, Y. et al., Investigation on the process of laser surface melting using two sequential scans. The International Journal of Advanced Manufacturing Technology, 2001, vol. 17, nos. 1 – 4, pp. 225 – 233. ISSN 0268-3768.

[5] PANDEY, A., DUBEY, A., Multiple quality optimization in laser cutting of difficult-to-laser-cut material using grey–fuzzy methodology. The International Journal of Advanced Manufacturing Technology, 2013, vol. 65, nos. 1–4, pp. 421–431. ISSN 0268-3768. doi: 10.1007/s00170-012-4181-5.

[6] GENNA, S., MENNA, E., RUBINO, G., TAGLIAFERRI, V., Experimental Investigation of Industrial Laser Cutting: The Effect of the Material Selection and the Process Parameters on the Kerf Quality. Applied Sciences, 2020, vol. 10, no. 14, article 4956. ISSN 2076-3417. DOI: 10.3390/app10144956.

[7] SHAMLOOEI, M., ZANON, G., BRUGNOLLI, M., VANIN, M., BURSI, O.S. (2024). Optimization of Surface Roughness in Laser Cutting Process of Mild Steel Using RSM and GA Method. In: SILVA, F.J.G., PEREIRA, A.B., CAMPILHO, R.D.S.G. (eds) Flexible Automation and Intelligent Manufacturing: Establishing Bridges for More Sustainable Manufacturing Systems. FAIM 2023. Cham: Springer, 2024, pp. 883–892, Lecture Notes in Mechanical Engineering, ISSN 2195-4356. doi: https://doi.org/10.1007/978-3-031-38241-3_91.

[8] TONTCHEV N., (2014) Materials Science, Effective solutions and Technological variants, Saarbrücken, Germany: LAP LAMBERT Academic Publishing, 2014. 144 p. ISBN: 978-3-659-57778-9.
[9] TONTCHEV, N., GAYDAROV, V., HRISTOV, N., Analyses and Applications of the Computational Approach "Define by Displacement" DEFMOT, Sofia, Todor Kableshkov University of Transport, 2022, 167 p. ISBN 978-954-12-0299-9.
[10] VUCHKOV, I., BOYADJIEVA, L., Quality Improvement with Design of Experiments. A Response Surface Approach, 2001, 508 p. ISSN 1566-0443. https://doi.org/10.1007/978-94-009-0009-7

✉ **Dr. Nikolay Hristov, Senior Lecturer**
Department of Technology Organization and Management of Transport
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: n_d_hristov@vtu.bg

✉ **Dr. Victor Georgiev**
Department of Technology Organization and Management of Transport
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: viktor_georgiev22@abv.bg

https://doi.org/10.53656/isct-2025.29

Architecture
and Civil Engineering

RESEARCH ON THE STABILITY OF CAISSON FOUNDATIONS

Stoyna Kostova
Todor Kableshkov University of Transport, Sofia, Bulgaria

Abstract. In the article general information for a brief overview of the problem status is given. Main principles of operation of caisson foundations and peculiarities in the design of deep foundations and in particular foundations with caissons are shown. Study of the stability of caisson foundations is one of the main tasks in their design. The article discusses the most significant problems related to the sizing, calculation and relevant inspections of these engineering facilities. A check of the bearing capacity of the ground, in specific engineering and geological conditions and loading from the superstructure are carried out. Technical measures in preparing the bottom of water basins for placing caissons, including materials used, specialized mechanization and process technology, ensuring the stability of the bottom of water basins are given. Special measures and regulations for the construction of the caisson, guaranteeing its stability are required. Features of transporting and laying caissons in open waters. Inspections of the deep foundations described in working and transport conditions. Computational checks and computational methods in accordance with Eurocode 7. Geotechnical. design. Part 1: Basic rules. BDS. EN 1997-1 [1, 2] are used. When solving problems related to the design of deep foundations / caissons /, the methods of Brinch - Hansen, Yanbu, Krey – Bishop etc. are presented in this article. The material is illustrated with an example and solution to a specific engineering problem. Based on the research, relevant conclusions and recommendations have been made.

Keywords: deep foundations; caisson foundations; stability check

1. Introduction

Caissons are used in underwater construction. Some of them are used to providing a working space when performing deep foundations. Excavation of the soil in the area in the working chamber was carried out by workers in the past. Today, excavation work is increasingly done mainly by machines with very little human presence. When performing this type of excavation work, it is necessary to protect the workers in the chamber from prolonged stay in it. To equalize the pressure in the circulatory system of people when working at depth at a higher pressure with the atmospheric pressure that is on the surface, pressure chambers are used. The danger lies in not getting the so-called caisson disease, which is not allowed in the current state of science and technology. Another type of caisson is of the “floating” type [3, 4, 5]. They are used for the implementation of foundation and enclosing structures performed underwater. Caissons are calculated for construction and operational conditions. It is necessary to specify the weight at which the caisson will float and the weight at which it will sink and assume its design (project) position Some of principles are shown in [6, 7, 8].

2. Existing conditions

The soils on site have been investigated to a depth of 33.10m. The terrain elevation is 2.22m. At a depth of -8.58m from the terrain (10.80m) height, there is a compacted blasted rock mass with bulk density in the natural condition γn =18,5 kN/m^3, φ=40°, cohesion c=0 kN/m^2, unit weight underwater γ' =11,0 kN/m^3. The rock pieces are sized from 0 to 200mm. At a depth of -8.58 m from the terrain (10.80 m) there is the same compacted blasted rock mass with dimensions of 0 to 200 mm.

The layer below to an elevation of -13.48 m (with a thickness of 4.90 m) is medium-sized silty sand, with small gravel and shells with the following indicators: γn =22.1 kN/m^3, angle of internal friction φ=24.4°, c=0 kN/m^2, γ' =12.1 kN/m^3. Under this layer we have a clayey layer with shells, and with a thickness of (1.7 m), γn =18.5 kN/m^3, φ=16.6°, c=16.4 kN/m^2.

Below it to a depth of -28.58 m and a thickness of 13.40 m lies swelling clay with limestone grains with small gravel with a conditional calculated soil load q_{R0}=300 kN/m^2. Under them there is a 1.5 m thick sandy clay with interlayers of sand with shells. And coarse clay sand with shells and gravel with a thickness of 0.8 m. The last studied layer is powder clay with γn =18.1 kN/m^3, φ=17.4°, c=16.9 kN/m^2, E_0= 261 kg/cm^2. Fig. 1 shows the location of the layers and their thicknesses, and physical and mechanical characteristics.

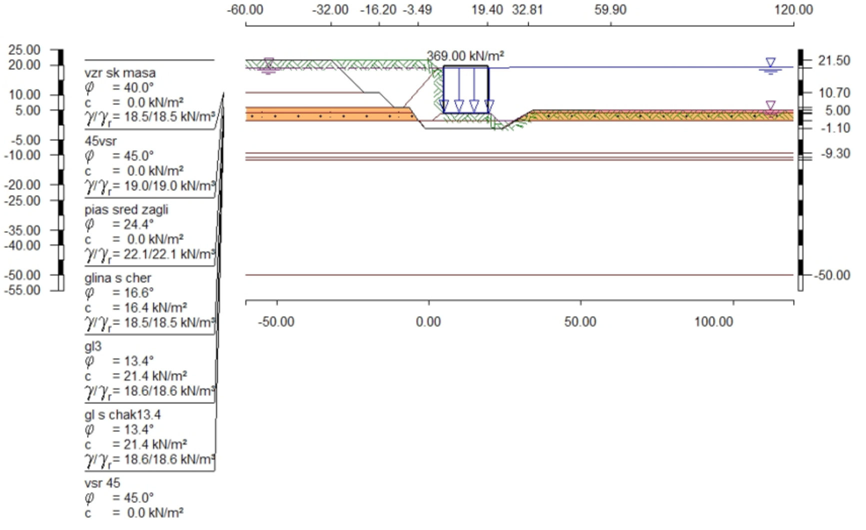


Figure 1. Location of the layers with thicknesses, physico-mechanical characteristics and load from the caisson

3. Technical measures for preparing the bottom of the water basins for placing caissons

The caisson is laid on two new layers with a thickness of 2.50 m. The first lower layer is made of blasted rock mass with the following physical and mechanical characteristics: γ' =11.0 kN/m^3, φ=40°, c=0 kN/m^2. This layer is laid on a previously prepared and aligned / dredged / base from the existing layer at the elevation of -19.28 m. At the depth of 21.40 m, the main layer is prepared, which is existing, swelling clay with limestone grains and with small gravel with a conditional calculated soil load q_{R0}=300 kN/m^2, unit weight in the natural condition γn =18.6 kN/m^3, angle of internal friction φ=13.4°, cohesion c=21.4 kN/m^2, filtration speed vs =425 m/s. The upper layer on which the caisson is placed has: γ' =12.0 kN/m^3, φ=45°, c=0 kN/m^2 and a weight of the stone pieces of 5 – 100 kg.

The two sub-layers of blasted rock mass are plased successively and then are compacted. Compaction can be done with vibrating plates. After compaction the rock mass embankment should be raised in accordance with the expected settlement of the caisson. After compaction, an excavator equipped with GPS sensors can be used for more precisely align the base of the footing. Divers can also be used for alignment, and they can inspect and assess the condition of the embankment on site.

To ensure the stability of the filled material, it is good to lay the blasted rock mass in a trench dug in advance for this purpose.

4. Safety checks when the caisson interacts with the earth's foundation

The caisson subject of the article is of the floating type. The caisson used for the expansion of the port of Burgas in Bulgaria has a prismatic shape. Its dimensions are 14.50 m wide, 14.00 m high and 60 m long. The reinforced concrete bottom plate and walls are reinforced and concreted in advance under special conditions. The reinforced concrete walls are executed up to about 6.00 m high. After which the walls are completed. The caisson is filled with water during construction and during its transportation to the project site. After its installation and consolidation, it is filled with sand.

The conditions of the caisson was investigated. The water level is at an elevation of -0.28 m and the caisson is completely below this level.

The load from the caisson has been calculated. A trapezoidal load with σ_{max} =540.0 kN/m^2 and σ_{min} =197.0 kN/m^2 has been obtained. For the calculations, the load has been averaged and assumed to be uniformly distributed with a value of 369.0 kN/m^2.

4.1. The check for the bearing capacity of the soil under foundation (Fig. 2)

This check was carried out according to Brinch Hansen and with Eurocode 7, Geotechnical. design. Part 1: Basic rules. EN 1997-1 [1]. It is done for 1 m in a perpendicular direction. Design value vertical loads Nd /Design value bearing capacity Resistance Rd is equal to 0.16. The check is fulfilled.

The solutions for the caisson were made with the DC Foundation computer program and its different modules DC Slope, DC Bearing, DC Settlement, DC Cantilever [9].

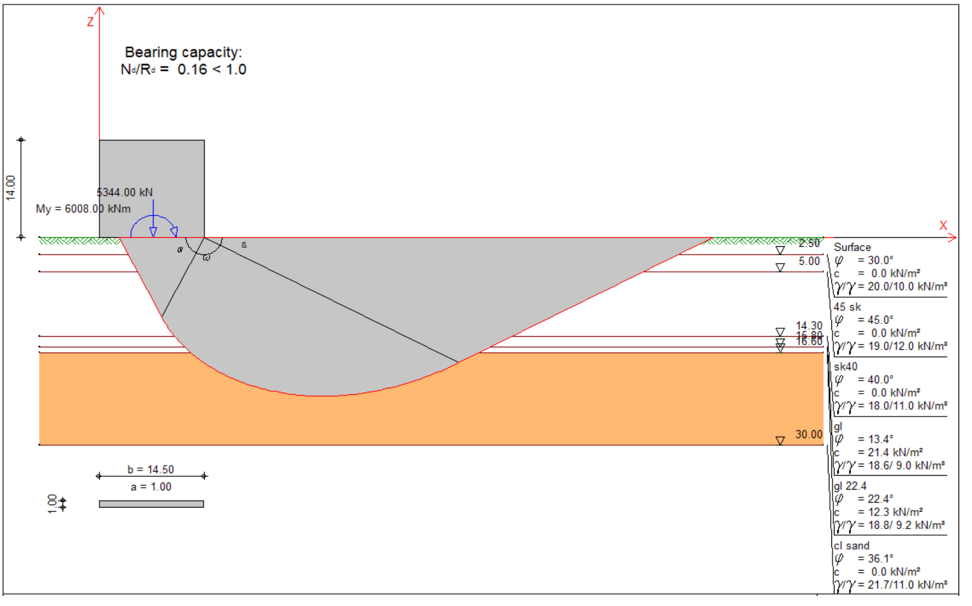


Figure 2. Bearing capacity of the soil base

4.2. Slope stability checks

First a check was made if a layer with a smaller angle of internal friction φ=40° under the caisson is placed. The stability of the caisson in this case is slightly above unity, which indicates that it is unstable. Factor of safety is 1.03 > 1.0. According to this check, it was decided that a layer with larger sizes of rock pieces and an internal friction angle of φ=45° should be placed under the caisson (Fig.3).

The caisson was checked in a state before backfilling on one side. In the operational state on the one side, the space next to the caisson is filled with blasted rock mass with the characteristics of the layer on which the caisson lays. After the calculations made according to Cray Bishop, a stable state of the caisson was obtained with a safety factor of 0.97 < 1.0.

After filling with a fraction of blasted rock mass with an angle of internal friction φ = 45° and stone fragments with a size of 5 to 100kg. The caisson has been checked for stability according to Krey – Bishop and acc.: DIN 1054:2005 EN [10]. According to this standard the Partial safety factors are given in Tabl.1.

Table 1. Partial safety factors (GEO) for LS 1C

γ	G	Q	W	E	φ	c	cu	R,a	R,c
Load comb. 1	1.00	1.30	1.00	1.00	1.25	1.25	1.25	1.10	1.40
Load comb. 2	1.00	1.20	1.00	1.00	1.15	1.15	1.15	1.10	1.30
Load comb. 3	1.00	1.00	1.00	1.00	1.10	1.10	1.10	1.10	1.20
Load comb. 2/3	1.00	1.10	1.00	1.00	1.13	1.13	1.13	1.10	1.25

The dates are: Raster with x from -7.90 m to 8.50 m, z from 22.60 m to 30.40 delta x = 5.00 m, delta z = 5.00 m,
with radius from R = 16.00 m to 31.00 m, delta R = 1.00 m; Load case k (Type: Lc 2)

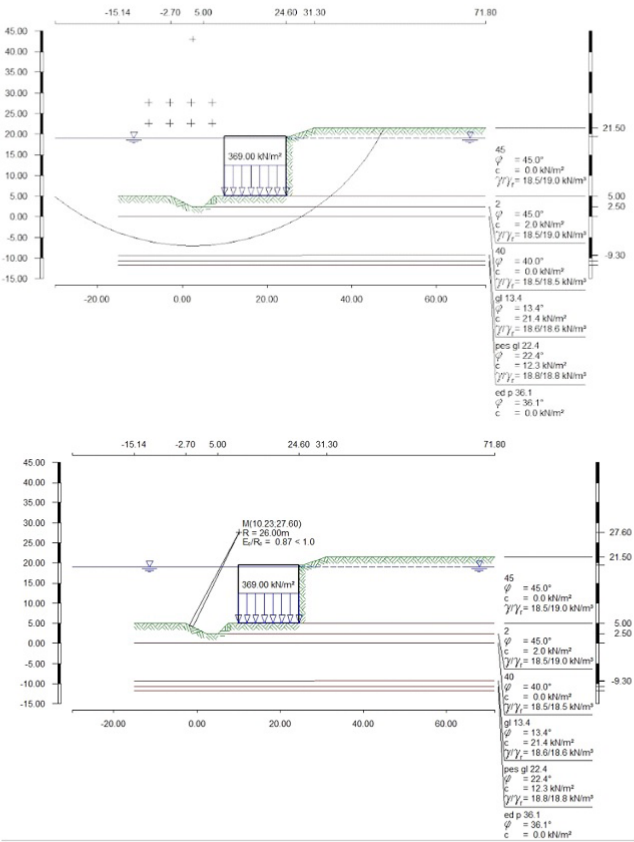


Figure 3. Slope stability check (Deep sliding check)

1. slope
Fail.plane from x = -1.89 to -0.91 m
Fail.circle:x M = 10.23 m, z M = 27.60 m, R = 26.00 m
2. slope
Fail.plane from x = 3.44 to 35.50 m
Fail.circle:x M = 10.23 m, z M = 27.60 m, R = 26.00 m
***** Section by a building
Loads Ed = 0.43 kN
Resistances Rd = 0.50 kN   Ed/Rd = 0.87 < 1.0, see Fig. 3.

Calculation according to British Standard EN 1997-1 – Design Approach 1 Combination 1 and 2 gives the security factor of 0.91 < 1. It is being shown in Fig. 4.

For Combination 1 with partial safety factors of groups A1 “+” M1 “+” R1

For Combination 2 with partial safety factors of groups A2 “+” M2 “+” R1 are used.

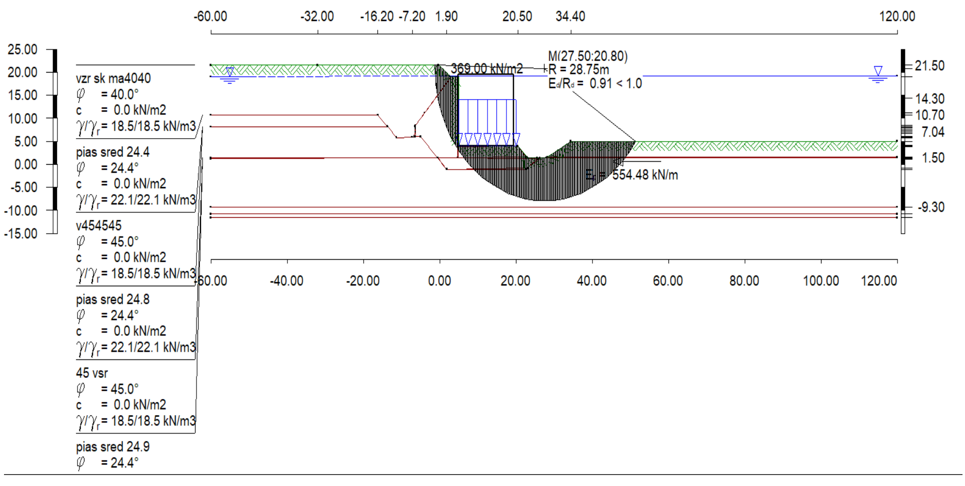


Figure 4. Slope stability check BS DA1

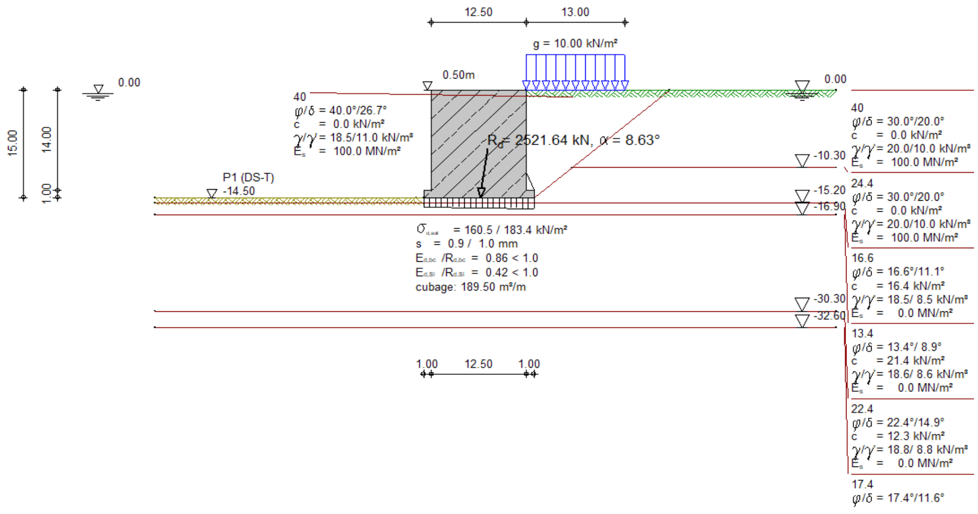


Figure 5. Check in operation conditions

4.3. Cantilever check (Fig.5)

The caisson was checked in working conditions. During its operation, the condition of the caisson approaches to the working of a supporting strengthening structure. At this stage, the caisson is placed in the project position.

It is filled with sand and its weight is determined, after that earth pressure diagrams were obtained. The decision was made according to Eurocode 7 BS 1997-1 [11], Design Approach 1 and combination 1 and 2 and their corresponding safety factors (Table 2). The water acts on the entire height of the caisson. It is completely submerged under water.

After the calculations Static equilibrium is fulfilled, sliding resistance is fulfilled 0.47 < 1.0. Bearing capacity is ok – Nd/Rd = 0.86 < 1.0. All checks of the reinforced concrete have been completed. All verifications are fulfilled.

Table 2. Partial safety factors for section forces (STR), for safety against sliding (GEO), for bearing capacity (GEO), for slope stability (GEO)

γ	G	E0g	W	L	E0l	Q	Qv	EP	Wg	φ	c	R,h	R,v
Comb. 1	1.35	1.35	1.35	1.35	1.35	1.50	1.50	1.00	1.00	1.00	1.00	1.00	1.00
Comb. 2	1.00	1.00	1.00	1.00	1.00	1.30	1.30	1.00	1.00	1.25	1.25	1.00	1.00

Conclusion

The calculation of caissons is very responsible and complex engineering work. The specific existing conditions at the bottom of the water basin must be considered during the design. According to these conditions and the careful work of the engineers, numerous checks, acceptances and decisions are carried out, which leads to the design of a stable and long-lasting facility.

From the calculations according to different European standards, several conclusions can be drawn.

When we use European standards, multiple solutions are obtained, corresponding to the accepted combinations and different partial safety factors.

It is good for complex and responsible facilities to carry out calculations according to more than one European standard, to cover all possible combinations of coefficients and ensure the necessary stability of the facilities during their operation.

REFERENCES

[1] EUROCODE 7, Geotechnical design. Part 1: Basic rules. EN 1997-1, Bulgarian Standardization institute.

[2] EUROCODE 7: Geotechnical design Part 1: Basic rules National Annex (NA), 2015, Bulgarian Standardization institute.

[3] https://www.bricknbolt.com/blogs-and-articles/construction-guide/floating-caisson.

[4] PANTOUVAKIS, JOHN-P., PANAS, A., Computer simulation and analysis framework for floating caisson construction operations, Automation in Construction, Volume 36, December 2013, Pages 196 – 207, ISSN 0926-5805, http://dx.doi.org/10.1016/j.autcon.2013.04.003.

[5] HERRERO, E., LLATA, J., G.SARABIA, E., VELASCO, F., SAINZ, J., R.LUIS, A., RUANO, S., GUANCHE, R., Dynamic positioning of floating caissons based on the UKF filter under external perturbances induced by waves, Ocean Engineering Volume 235, 1 September 2021, 109055, ISSN 0029-8018, https://doi.org/10.1016/j.oceaneng.2021.109055.

[6] XU, C., JIANG, H., XU, M., SUN D., RUI, S., Calculation Method for Uplift Capacity of Suction Caisson in Sand Considering Different Drainage Conditions, Sustainability 2023, 15, 454, https://doi.org/10.3390/su15010454.

[7] DENG, W., CARTER, J., A Theoretical Study of the Vertical Uplift Capacity of Suction Caissons, International Journal of Offshore and Polar Engineering, Vol. 12, 2002, pp. 89 – 97, ISSN 1053-5381.

[8] DENG, W., CARTER, J., Inclined Uplift Capacity of Suction Caissons in Sand, May 2000, OTC 12196, pp. 1 – 12, doi:10.4043/12196-MS.

[9] https://www.dc-software.com/home.

[10] DIN 1054:2005 Subsoil – Verification of the safety of earthworks and foundations, by Deutsches Institut Fur Normung /German Standards/ 123 p.

[11] BRITISH STANDARD BS EN 1997-1: Geotechnical design. Part 1: Basic rules.

✉ **Dr. Stoyna Kostova, Assoc. Prof.**
ORCID iD: 0009-0001-4365-0690
Department of Transport Construction and Equipment
Todor Kableshkov University of Transport
158, Geo Milev St., 1574 Sofia, Bulgaria
E-mail: skostova@vtu.bg, kostova.stoyna@gmail.com